PhD
1930
P

Ph.D. 1930 P
N.E. Powers
c.1

Boston University
College of Liberal Arts
Library

THE GIFT OF The Author

PhD 1930
P
c1

June 1930

44386

MEASUREMENT OF INTELLIGENCE OF SCHOOL CHILDREN AT WORK

PROBLEMS ARISING
during
A THREE-YEAR STUDY IN THE SCHOOLS
of
BELMONT, MASSACHUSETTS

by

Nellie Eva Powers

(B.S. in Education, Boston University, 1923; A.M., 1925)

A Dissertation
submitted in partial fulfilment of the
requirements for the
degree of Doctor of Philosophy

GRADUATE SCHOOL

BOSTON UNIVERSITY

1930

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

44386

PhD
1930
P
copy 1

PREFACE

Three years of study of a typical situation have shown certain problems as insistent and significant, during the planning and execution of a test program for the classification of the pupils of a school system. No claim is made that the most significant problems have been located. Attempt has been made to better understand the situation through use of the best tools available; and in the light of the data to determine what may best be done to obtain a working classification. It is hoped that the record of the problems encountered, and of the work done in attempts at solution, may reduce the blunders of those who undertake similar problems elsewhere.

Grateful acknowledgment is made to Superintendent Frank A. Scott and to all of his administrative and teaching force for their labor and coöperation; to the Summer Session Class of 1927 at Boston University in Measurement of Intelligence, for assistance with interpretation of data; to all of the authors whose works are quoted herein; to the sixty-one heads of research departments in public schools and other leaders in the measurement field who have contributed materials and answers to the nation-wide questionnaire; and especially to Professor Guy M. Wilson for the privilege of his guidance and inspiration.

Nellie Eva Powers

Boston, April 30, 1930.

CONTENTS

Chapter Page

I. Should schools make adjustments for individual differences? If so, how can they do it? 1
Do individual differences exist, and are they discoverable? 1
Are the found differences matter of concern? 4
Are the old educational philosophy and practice, which took no note of child abilities, adequate to meet today's call for adjustment to individual differences? 5
What is intelligence, and can it be measured? 7
Does ability grouping help pupils and teachers? 10
Is there opportunity for blundering? 11
Is blundering avoidable? What is a workable plan? How much does it cost? How much does it save? 13

II. Purpose and Basis of the Study........... 14
Main problem: Can a program of homogeneous grouping be accomplished with the coöperation of all concerned, and with benefit to children and teachers? 14
Secondary problem: 14
(a) What problems arise in a program looking towards homogeneous grouping of pupils in a school system? 14
(b) What is the cost of such a program in time, labor, and money? What are the compensations? 14
Why haphazard testing is of little value. 15
The goals at which to aim............... 17
Basis of the study: The school system of Belmont, Massachusetts........... 17

III. Securing coöperation and understanding in the development of a testing program............................ 22
What general plans can be worked out? ... 22
What tests shall be given? 26
How can the teachers be brought into the plan? 33

Chapter Page

IV. How rapidly and how can teachers be trained to administer group intelligence tests? 37
What problems arise in the administration and early development of a program? 37
What are the problem outcomes of the first testing? 40
Problem a. How can the data be returned to the teacher in order to be of immediate service? 40
Problem b. What office recordof test results is needed? 46
Problem c. Given the data from a test program, how can classifications be made? 49

V. How rapidly can all teachers and pupils of a school system be included in the program of testing and classification? 56
What problems follow the first classification? 61
Where is the best place to begin intensive work, and what can be done? 61

VI. Developing the need for assistance....... 68
How rapidly can consciousness of needs be developed? 68
What are the possible and best plans for extra assistance? 69
How can the expense of extra help be justified? 70
What are the new problems of a new worker? 74

VII. How can the entire organization be involved in work upon the problems of classification of pupils? 77
Can blunders be avoided? 77
Can coöperation be secured?.............. 78
Can the entire organization be involved? If so, how and at what cost?........ 82

VIII. What are the compensations for the expense of time, labor, and money? 86

Chapter Page

IX. Supporting data.............................. 95
Nation-wide questionnaire................. 99
Leaders in public school research........ 107
Answers to questionnaire.................. 108

X. Summary.................................. 132
Re-statement of problem.................. 133
How the evidence was obtained............ 133
What was found........................... 135
Answers to problems which arose..... 136
Two sides of the ledger............. 137
A program for success.................... 139

Bibliography.. 140

Autobiography... 151

TABLES AND CHARTS

Table

Page

1 School enrollments in Belmont............ 18
2 Preparation and experience of teacher personnel in Belmont............... 19
3 Dougherty's Correlations of certain tests with Stanford Revision of the Binet Tests..................... 30
4 Test data of three typical cases......... 52
5 Ability, health, conduct and scholarship of six superior and very superior pupils of Senior High School.............................. 60
6 Records of 18 High School pupils who have repeated at least one grade during their school careers... 60
7 Promotion data, June, 1927............... 62
8 Years of repeating done by pupils enrolled in the Belmont schools..... 71
9a Cost to Belmont of Retardation........... 72
9b Savings for Belmont from Acceleration.... 74
10 Cost of Testing Program for three years.. 84
11 Hours of Labor and Teaching given by the organizer during three years of preliminary work.................. 85
12 Leaders of Public School Research, with cities and populations which they serve........................... 107

Chart

I IQ's of pupils of Grade II as shown by Pintner-Cunningham Primary Mental Tests........................ 43
II IQ's of pupils of Grade II as shown by Detroit First Grade Test......... 44
III IQ's of pupils of Grade II as shown by Haggerty Intelligence Examination, Delta 1....................... 45
IV IQ's of 232 kindergarten pupils.......... 65
V Graph of years repeated by the school population of Belmont........ 73

CHAPTER I

SHOULD THE SCHOOLS MAKE ADJUSTMENTS FOR INDIVIDUAL DIFFERENCES? IF SO, HOW CAN THEY DO IT?

Do individual differences exist and are they discoverable?

That certain men are taller than others, that certain men can run faster or lift heavier burdens than others raises no question as to the existence or validity of difference. Physical differences are for all practical purposes easily measurable in terms of universally accepted units, height, speed, weight, etc. The existence of these measurable differences of individuals in the physical field is popularly accepted, and this fact was utilized notably during the World War in making adaptations of worker to task.

The existence of individual differences in the mental field has come to recognition much more slowly. The source of mentality is hidden, making it necessary to evolve means of measuring mentality through its products in behavior. Up to very recent times teachers have felt that one child could learn just as well as another "if he would only put his mind on his work", or "if the teacher made him do it". With all of the "putting of mind on work", and with all of the enormous teacher effort, chil-

dren of twelve to fourteen years of age were to be found in the first three grades, large numbers left school as soon as possible, and other large numbers never entered the High School.

Much early work in development of tests was founded upon physical measurements, and sought likenesses rather than differences. (23: 373-380)[1]. Up to 1896 when Cattell made his first report of measurement work at Columbia University (24), psychological thought and educational practice were in terms of "the mind" which needed discipline and of"faculties of the mind" which needed training. Binet's signal contribution of a scale for measuring intelligence (13,14,17,15) brought a new concept of intelligence and of the basic qualities of an intelligence test. He strove to measure judgment, common sense, initiative, the ability to adapt one's self to circumstances and the ability to accomplish unit mental tasks (16,12). Early work with Binet tests in America,-- that of Goddard (40,41,42), Huey (49), Kuhlman (58,59), Melville (67), Terman (93,85,92), and that of Yerkes and Bridges,-- adapted the scale and increased interest in the results of its application. The use of mental tests in the army, and of group tests in schools has made a large scale study possible.

---

[1] The first number within the brackets refers to the book or article in the Bibliography. Numbers following the colon are page numbers.

Terman shows that when unselected groups of children are examined for any trait, physical or mental, large differences are found; and if the groups are large enough the differences are distributed along the normal curve of distribution. (86: Ch. II) Nosworthy, in her study of inmates of institutions for the feeble-minded, found there normal distribution of height, weight, intelligence, etc., though the intelligence was at the low end of the scale. She also showed, from taking a vertical sampling at a given point in the distribution that there is much overlap. There is no hard and fast line between, for example, normality and moronity. She shows that, in certain physical characteristics, 48% of the inmates of the feeble-minded institutions studied, are on a par with their normal compeers. (69) Thus it appears that mental traits are distributed on a normal curve, as are physical. In short, the normal distribution is nature's method, and no exception is made in mental traits.

Terman and his associates give data which show that "from 20 to 25 per cent of the pupils of a given grade have attained a mental age about as high as the median mental age of the next higher grade, while the lowest 20 to 25 per cent in the same grade are about as low in mental age as the median for the next grade below." (94: 8-9)

It is more and more evident to all that individual

differences in mental ability is a demonstrable fact, and that measures of those differences yield data for classification, point out the slow and the gifted, and assist in making adjustments in curricula and methods.

Are the found differences matter of concern?

Terman again says, "It is doubtful whether in an average school system more than 70 per cent of the pupils are being given instruction which is as well suited to them as would be the instruction in a higher or lower grade". (94: 12) In his study of the gifted, reported in Genetic Studies of Genius, "82 per cent of the cases are said to merit some additional promotion". (89: 285) Wentworth from her study of eleven superior children, shows the great variations in the abilities of the gifted, and the fact that "because teachers were trying to help children to reach impossible standards", much of the teaching of the slow was lost. (105: 84-95) Mort pictures a fourth grade group in which "more than half the pupils are overage for their grade, one girl being of junior-high-school age. The range in reading age is from 5 years, 10 months, to 13 years, 8 months, and shows need of individual assistance. This situation is not unusual". (68: 265) These findings offer challenge to the efficiency of our schools; and as such, the differences are matters of serious concern. Undoubtedly Terman would state the above differently today. Rapid progress is being made

in thought and work upon the problems of individual differences. This is shown forcibly in Broady's study of school provision for individual differences in eighteen cities of the United States. (19) Children should be grouped on the basis of ability and given sizable tasks for age, regardless of the old grade scheme.

Are the old educational philosophy and practice, which took no note of child abilities, adequate to meet to-day's call for adjustment to individual differences?

The development of the theory of individual differences and modern attempts to make adjustments in the schools which shall be commensurate with the found differences, have come hand in hand with the changing times. Dickson expresses the cause and the change, thus: "The period of compulsory education, by forcing all types of children into the public schools, brought about still greater need for classification, and for variation in the curriculum. These factors are involved in bringing about the present conception of the aims of education; i.e. that every child, of whatever capacity or probable destiny, shall be given such opportunity for development as will enable him to realize his best possibilities and satisfy the requirements of good citizenship." (33: 8-9) The old school as a static instructional unit is growing towards the new school as a living place for individuals who are learning how best to live. The times have forced

the facts upon us. Measurement, especially measurement of intelligence, has given us a tool, the right use of which can point the way for adjustments in philosophy and practice, to the newer needs. Shall we call it conservatism, or stupidity, when educators cling to the last straw of the old, rather than learn to swim in the great, new, open sea? Such shift in aims of education, such adjustments in philosophy and practice could hardly be expected to come about without arousing popular interest and discussion, and a certain amount of questioning.

The intelligence test movement which sought to find differences pointing the way to better adaptations, has experienced a generous share of popular questioning. Popular discussion based upon unmodified and uninterpreted catch phrases from reports of the army testing (113,18), appeared in newspaper and periodical (116,35,21,61,28). It was unfortunate that Bagley chose to cling to the old, and to kindle discussion based upon the emotional appeal of the newspaper and the popular magazine (8,9,10). Individual differences became an excuse for lugging democracy into the discussion. Measurement workers, led by Terman and Whipple, showed that misstatement and misunderstanding prevailed, and that there was need for careful definition, for outline of purpose, and for accuracy of work and report in the field (87,88,106,107). Even Kelley tried to balance up "argyments", and call the results "salvage" (54: 11).

The problems of individual differences are present in the schools, and can be solved not by sitting down and talking and thinking it over, but by a rigid following of the technique of research, and an open-minded acceptance of the findings of research. Philosophy and practices cannot continue as in the "good old days". The controversial blunders of some educational leaders have come about through willingness to allow "say so" and "thought about it" to dictate, rather than data and deducible facts from data. The appeal has been to prejudice instead of to right interpretation of data. Thompson succinctly summarized the situation in 1924 as the suffering of the general intelligence and educational test movement from growing pains, and from the conservatism of the human mind towards all things new (95: 5). Johnson makes prediction that "the use of measurement with regard to certain aspects of human nature is going to be increasingly helpful as we learn more about methods of measurement and content of human nature". (52: 5). To-day's call is for a "child first" philosophy and practice. The child, his abilities, his idiosyncrasies, and needs must indicate the course of adaptations. Philosophy and practice must renew or adjust.

What is intelligence, and can it be measured?

Early in 1921 an attempt was made by the measurers to define intelligence (99). All of the definitions stated or implied that intelligence is recognized as a composite

whose nature still demands intensive research. Thorndike offered as a working definition: "Ability to deal with things or persons or ideas by the use of ideas" (98: 413). He attempted to make this more specific by saying that we must realize that "ideas are not like eggs in a basket, nor like eggs in an omelet, but are what they are. If we must liken them to something, let it be tones in a symphony or factions in a party, or neurones in the brain or some other case of a very complex organization where science can and should separate the totals into parts, but where those parts are splitting and combining from time to time and are being influenced in their action by more or less of the total organization" (98: 419).

In an attempt to clarify the meaning and nature of intelligence Spearman worked out his Tetrad Difference Formula by means of which he proved the existence of G (a general ability factor affecting all mental life). He was able to separate out other specific factors (79: 136-160). Kelley has gone a step further than Spearman, and by means of his Pentad Criterion claims to have been able to break up Spearman's G factor into general and other independent factors (55: 34-96). Spearman sees "surprising agreement in the two differing techniques", and "extraordinarily harmonious results" (80: 568).

Binet and Simon have said: "On peut mesurer l'intelligence", disent les uns. "L'intelligence ne se mesure

pas", disent les autres. D'autres, mieux avises, dedoignent ces discussions theoriques, et s'appliquent a resoudre le probleme en fait (17: 1). (Some say that intelligence can be measured. Others say that intelligence is not to be measured. Still others, the better advised, disdain these theoretical discussions and apply themselves to solving the problem by seeking the facts.) The same situation exists to-day with regard to intelligence tests. Testers see the values in the use of tests; a dwindling number of observers and "thinkers about" may be skeptical. The facts must be found. Bronner, Healy, Lowe and Shimberg consider that "mental testing requires no justification; the rapid growth of belief in its value has in many ways proved to be well warranted" (20: 3). "It is true that many people oppose these tests; but the fact is that there is practically no competent educational psychologist to-day who repudiates mental tests..... This should outweigh any non-experimental objections to the tests" (118: 273). Terman makes the prediction that a hundred years from now the educational historian will probably characterize the present era as the one which saw the birth and development of the testing movement. From a half dozen points of view the introduction of standard tests of intelligence and of school achievement must be regarded as of capital importance, for these devices are already leaving their mark upon school organization and

administration, instructional method, curriculum content, text books, and even educational ideals (90: 57 also 45: 260).

Does ability grouping help pupils and teachers?

The question is answered thus by the workers quoted:

Dickson: "Pupils of similar ability work better together, and the larger the classes which can be handled the lower the cost of education will be" (94: 32).

Terman: "A reasonable homogeneity in the mental ability of pupils who are instructed together is a sine qua non of school efficiency". (94:7)

Allen: From right grouping "many of these pupils (those who never go to High School) may be so educated that their earning power can be doubled" (2: 13).

Corning's classification "held superior pupils to a higher standard than that to which they were accustomed, and all so placed finally succeeded at the higher level" (29: 165).

Torgerson: "Proper classification and sectioning reduces failures and tends to raise the accomplishment quotient of all pupils to maximal efficiency" (100: 180).

Worlton: Under ability grouping we find "more rapid progress in subjects; more congenial social atmosphere in which to work; more opportunity to develop leadership; greater stimulus for individual effort; increased happi-

ness from successful accomplishments; better opportunity to curb snobbishness; fewer disciplinary problems". "There is also better morale among teachers" (112: 274).

Frazee: "In its last analysis, ability classification rests upon the bed-rock of human relationships. Discouragement and conceit have little soil in which to grow. Understanding and tolerance develop. Each individual measures himself against his own group and so gets a truer estimate of himself. On each level and sub-level leadership is bred" (37: 77).

Maller: "Members of a group will tend to show a higher degree of coöperation if there is, within the group, little variance in intelligence, age, and social factors. A maximum of coöperation is obtained when a group is of highest homogeneity. The sectioning of classrooms on the basis of intelligence and other factors would thus augment the cultivation of coöperation" (63: 163).

Is there opportunity for blundering?

In the town in which this study was made blunders had accumulated until the superintendent himself took courses in measurement and sought advice at Schools of Education as to how the blunders might be righted. A serious mistake is made when testing, a tool for use in reaearch, is reckoned a desirable new "frill". As though testing could, from its very nature, have a "frill", much more be one! In this town there was no recognition of the need

for careful, definite, skilful planning for a coördinated program. Aims, goals, purposes of program, knowledge and understanding of tests and testing were worse than nil, for there was no recognition of need for them. The tests were purchased, distributed and used as best they might be. The time for giving, scoring, and returning the first tests was ignored. The testing was, accordingly, sporadic, because of lack of knowledge and leadership and guidance. Furthermore, teachers were asked to give the tests and to judge them, when they had little, if any, basis in knowledge for judging. Later, teachers were asked to keep a record of the time consumed in scoring Stanford-Achievement Tests, when they had no knowledge of methods of arranging nor handling papers for scoring, and when they were much puzzled over details from test manuals, and wearied with the conscientious effort which they made to do their best. Procrastination, lack of sympathy, and biased critical attitude towards tests, was developing by the very testing as it proceeded. The need for expert help and guidance at the beginning was not felt. Willing and capable teacher-minds were permitted and encouraged to function upon a closed basis of opinion, rather than upon an open-minded attitude pointing toward the scientific.

Competent psychologists and the best workers in scientifically conducted research are united in seeing great values in the use of intelligence tests for securing

grouping (118: 273). Blunders in current criticism of such use of intelligence tests are based upon the willingness of the popular critic to use exact figures without the knowledge of nor consideration for the modifications and interpretations which research shows as vital.

Is blundering avoidable with properly trained teachers? What is a workable plan? How much does it cost? How much does it save? Answers must be sought in the data of those of experience, and in actual research.

This study was undertaken in an effort to undo the effects of past blunders in test matters, by training teachers to better understanding of their own problems in the field of individual differences through actual work in a coördinated program in classification. Tentative Answers have been found to these last questions.

CHAPTER II

PURPOSE AND BASIS OF THE STUDY

The main problem of the study is: Can a program of homogeneous grouping be accomplished with coöperation of and known benefits to teachers, with coöperation and understanding by parents, with benefits to and without detrimental effects upon children?

Secondary problems are:

(a) What problems arise in a program looking towards homogeneous grouping of pupils in a school system?

(b) What is the cost of such a program in time, labor, and money? What are the compensations for the cost?

Public school superintendents will vouch for the fact that many attempts have been made to use achievement and intelligence tests within their school systems from which there was little return for a vast amount of labor and cost. The data from such tests are gathering dust while awaiting interpretation. Tests have been given, as they came upon the market or became the fashion, only to add their contribution as dust gatherers, or to be laid aside and forgotten before reaching the central office. The most valuable information obtainable from testing is usually neglected. Teachers give tests, and after they have served the needs of the moment they are either thrown

away immediately or filed and forgotten. Not half the use is made of the responses of pupils to these tests as could and should be made (22: 341-342).

There has been misunderstanding of the purpose of testing (110: 319), blundering in the leadership or directorship of the testing, blundering in the administration of tests, and possibly most serious of all, blundering in the interpretation of data from tests.

Haphazard testing has been of little value in that its purposes and objectives were not carefully defined and understood; the testing was not planned nor worked out in detail in the light of the best experience and information available, and was not expertly directed. The testing has been more or less of a superimposition upon teacher and child. This study raises the question of whether such blunders are avoidable with proper direction, proper precautions, and proper training. An attempt is made to recognize significant problems which do arise, and to show how these problems have been anticipated and met in a given typical situation.

Intelligence test data are not as yet considered a safe, single criterion by which to form homogeneous groups (109: 262). The intelligence test is a tool which helps in the recognition of individual differences. Grouping of classes is not homogeneous until as many factors as possible have been taken into consideration. Health,

intelligence, academic achievements of the past and present, social backgrounds, manual and other skills, behaviors, all have their contributions to make. Intelligence test results are only one of many factors, but these results "are of value in classifying pupils in accordance with their natural intelligence, or ability to learn", says the National Research Council (115: 27). Dickson writes that "Evidence thus far presented has emphasized the mental age as the most important single criterion in determining the grade location of a child in school. ------- Such mental age has been determined by intelligence test". (32: 30, Sec. 4). Thorndike reports that "hour for hour, or dollar for dollar, the psychological test for intellect is preferable to conventional tests for scholarship. ---- The psychological test measures ability more, and promise more, and schooling less, than achievement tests" (96: 102). It was the experience of Fordyce in a Lincoln School Study that "correlations between intelligence test results and teacher estimates were sufficiently high to indicate that it is possible for one to discover by means of mental tests along with the grades and estimates of teachers such capacities as are essential in determining the proper grading and promotion of pupils". (36: 43). Accordingly, the beginnings of homogeneous groupings are assumed to be safely made from data carefully obtained from intelligence tests, this data being supple-

mented by health records, teachers' marks, and any other available authentic records of achievement or behavior. The major goal of a testing program has been the discovery, in the most objective way possible, of information which will lead to better understanding of the individual child, and thence to better adaptations to his found needs. Such a program demands skilled leadership, careful and thorough training of those who administer the tests, wise collation and interpretation of data, thorough understandings of meanings on the part of all concerned. There must be willingness to accept the facts as found, and to act promptly in accordance with these facts. This is not a matter of purchase of test material, experimenting with a test, and boasting of having a new frill in the system. It is a matter of careful research, skilfully planned work, and problem thinking on the part of all.

The study leads to the finding of values of an intelligence test program, in terms of expense in time, labor, and money, as checked against compensations.

The basis of the study:

The school system of Belmont, Massachusetts, furnished a typical situation, and was made available for study from September, 1926, to June, 1929. Belmont is a suburb of Boston, a rapidly-growing residential town with a population of 18,000. The population consists of two leading elements, the well-to-do American-born, and the progressive

Italian. There is a negligible number of other nationalities. Table 1 shows the number of pupils in the schools during the three years of the study.

Table 1

School Enrollments in Belmont

Number of Schools	Schools	Roll In 1926	Roll In 1927	Roll In 1928
1	Senior High School	606	693	790
1	Junior High School	552	550	591
6	Grade Schools	2,166	2,229	2,387
	Totals	3,323	3,472	3,768

There is an unusually large shift in school population each year. For the year 1927-28 the Superintendent reported "33-1/3% of the pupils in the elementary grades moved away, and a larger number filled their places". (Unpublished school data.)

The Elementary Schools contain kindergartens and grades through VI; the Junior High School serves seventh and eighth year pupils and a few ninth from an overflow group at Senior High School; the Senior High School serves ninth, tenth, eleventh, and twelfth year pupils, with a few thirteenth year pupils who are taking special work. Academic backgrounds of the teachers are shown in Table 2.

Table 2

Preparation and Experience of Teacher Personnel for the year 1927-28 in Belmont

Schools	Number of Teachers	Teacher Training	Years of Service
1 Senior High	29	Graduate of	Range 1-37
		University - 13	
		" of College ---- 12	Median - 3
		" of Normal School ----- 3	
		" of Academy ---- 1	
1 Junior High	24	Graduate of	
		University - 6	Range 1-18
		" of College ---- 5	Median 2.5
		" of Normal School ----- 12	
		" of Trade School ---- 1	
6 Elementary	65	Graduate of	
		University - 2	Range 1-39
		" of College ---- 2	Median - 3
		" of Normal School ----- 50	
		" of Rg. Training School -- 11	

These factors influenced problems of the experiment: previous tests and organizer's time. For five years previous to the experiment there had been sporadic attempts at measuring in these schools. The tests, as recorded by the Superintendent, had included: Stanford Tests in Achievement, 1925, Various High School Subject Tests,

Dearborn test, 1921-1926, Wilson Survey Tests in Arithmetic, Courtis Standard Tests in Arithmetic, Spelling Tests based on Ayres' Spelling Lists, Test on Jones' Demons, Kansas State Normal Tests, Burgess Silent Reading Tests, Ayres' Reading Test, Munroe's Silent Reading Test, Springfield Tests, Downey's Provisional Entrance Tests, Brown Psychological Test, Terman Group Test, and Ayres' Handwriting Scale.

Practically no use was made of test findings; often returns were left incompleted; the little material from these tests, which was available in September, 1926, was so erroneously scored and summarized as to be valueless. Early in 1926, 2000 Stanford Achievement Tests were attempted. The teachers, with little help, had found scoring difficult and time-consuming, and the results, arrived at in different ways, of questionable value. Material was not all collected.

Superintendent Scott, in an unpublished thesis written at Harvard Graduate School of Education in 1926 on "A Testing Program Administered by Teachers", questioned:

"1. What is the best way of getting test work done in time of service?

2. Does the giving and correcting of tests by teachers and principals seriously interfere with their use?

3. What administrative uses can be made of the tests when completed?

4. What advantages has teacher-administration of tests, over the giving by a person whose sole duty it is to give tests?

5. Does the attitude of the teacher bear an important ratio to the value of results?

6. Can teachers learn to substitute Standardised tests for tests of their own making?

7. Would a secretary for each building help solve the problems of correcting and tabulating?

8. Do inaccuracies of teachers mar the work?"

He also asked: "Why not use the old test data? We have so much of it which is doing us no good."

It was possible for the author, known throughout this dissertation as organizer, to serve for only part time for two of the three years of the study. This increased clerical work for organizer and teachers, added to the cost of postage, and lessened slightly the necessary personal contact of organizer and the schools. It did help to stress the need for a full-time director of tests and measurements.

The program was planned for the purpose of arrnaging better classification of pupils, if possible, than existed under the old age-grade scheme. In order that results might function in better adaptations, it was necessary to substitute a more wholesome attitude of teachers towards tests and testing. This was to be attempted through detailed training of the teachers to administer tests to their pupils and to understand and use results.

CHAPTER III

SECURING COOPERATION AND UNDERSTANDING IN THE DEVELOPMENT OF A TESTING PROGRAM

What general plans can be worked out?

The making of a workable plan presupposes the fact that a competent leader, responsible to the superintendent, is in charge. He is a person who not only knows and is experienced in research, but is one who understands that he is in charge of fact finding for the school system, fact finding which may shape policy but which will not make it nor administer it (78: 11), (26: 39).

Enough must be known of past testing and test experiences to make future plans such as to prevent recurrence of waste. For example: if, as in the typical situation, teacher attitude has been seriously warped in working long hours at scoring tests without ample direction and assistance, and with testing itself as the goal, then knowledge of these facts will help in planning a new program in such a way as to look towards counteraction of the old wasteful methods. Hughes says, "The success of the research will depend, in large measure, on the ability of the director to secure the coöperation of individuals and groups of individuals in an attempt to solve ----- problems" (50: 13).

A general survey of intelligence may be instituted in order to discover present conditions as shown by the tests used, and from the data thus acquired to obtain material which will determine future action. Objectives must be presented in such a way that the workers will not lose sight of larger purposes as the work goes on upon the more or less mechanical details of the test program. It is the consensus of reports of leading heads of research departments (IX: Q3, a and b)[2] that it is worth while for all concerned to attend a general meeting at which objectives and purposes of a plan are carefully defined. The giving of these objectives in mimeographed form may help to fix teacher attention upon them. In Cleveland, the mimeographed form, worked out in detail, suffices (117).

The initial plan which was carried out in Belmont was for teacher administration of intelligence tests to all pupils in the schools below Junior High School grade. The objectives of the study, laid before all of the teachers concerned, by the organizer at a general meeting were these:

---

[2] "IX" refers to the chapter in which answers to the nation-wide questionnaire are given.

"Q3, a and b" refer to the a and b questions under question 3.

(1) Health, physical, mental, and social, for each individual pupil (as far as it is possible to obtain it).

(2) Classification through specialized placing of the individual child. This may

Reduce retardation;
Reduce costs due to retardation;
Prevent non-promotion;
Make teacher work easier;
Make for happiness of teacher and child;
Democratize the school.

(3) Closer contact of the individuals of all departments with the problems of the individual child.

(4) Greater happiness of child and teacher from understanding and applying the "child first" principle.

(5) Scientific attitude of all.

Teachers need help in understanding each of these points. It is economy to arrange for ample time for explanation and teacher discussion of the points.

Since the testing was to be administered by teachers, the training of the teachers to give, score, and summarize the particular tests to be used, became a necessary part of the general plan. It was assumed that it would be possible to give this training most satisfactorily on the actual tasks on hand. Pintner shows two methods of training graduate students to use group tests: one, a study of the tests in logical order, with discussion, with some practice, and with attempt to evaluate results; the other, actual testing on a real situation. He finds that "Whatever may have been lost in logical presentation and

completeness of information about all the different group tests was compensated for by actual practice in giving tests in real school situations, and by scoring tests and tabulating results that were going to be used by the schools". (74: 272).

The amount of clerical work to be done in connection with the survey, and the adjustment of the clerical load of a school system to include the added amount or the employment of additional service, are also necessary parts of the general plan. For estimates of amounts of clerical service required see Table 10 (IX: Q7). Added clerical service includes correspondence, scoring and rescoring of tests, summarizing and tabulating of data, making of classification lists and extra promotion lists, card filing of data, and making of charts, graphs, and reports.

Size of school system, and size and organization of research personnel has much bearing upon the method of detailed work upon a test program (IX: Q5). When a survey is instituted the leader plans, or delegates the planning, of what tests are to be given, who is to give them, how and when they are to be given; sets the time and place for centralization of returns; and arranges other details necessary to the proper staging of the program and collection of data.

Very early, plans are needed for record forms and

files which will protect the data and expedite its later use, especially use for classification. Making the files easily accessible to the users is of importance.

Beyond the planning of a survey little can be done until survey results are available for study and interpretation. Thorough study and careful interpretation of the data yield material which outlines classifications and indicates the needs for future plans and work.

What tests shall be given?

In a study made by Root it is assumed that "Binet tests constitute the truest estimate of intelligence in so far as tests go. This may not always be the case with older (college) students, but probably no one will take exception to the assumption that with children up to fifteen or sixteen years of age, the Binet test constitutes the best single test estimate of the intelligence that can be made" (76: 286). The administration of Binet tests to large numbers is expensive in time, and in skill necessary for the examiner. Group intelligence tests offer means of testing large numbers at a considerable saving of time, with less demand upon the examiner. The question then becomes "What group tests shall be used?

Dickson gives the essentials of a good group test as:

It must "contain a large variety of mental problems".

It must "show a graduated scale of difficulty in each type of exercise".

It must "provide sufficient opportunity for differentiation between lower and higher mentality, so that the different grades of subjects, from the very superior through the average group to the very inferior, will each be separated from the others rather definitely in the final scores obtained". Such standardization should be made through the analysis of data from thousands of cases.

"The procedure for giving and scoring the test must be mechanical, permitting no element of personality to enter."

"There must be some provision made whereby, if the test is given during successive years, it is freed from the element of coaching."

It must be "so arranged that it can be given within reasonable limits of time" (33: 43-44).

Irwin and Marks utilize "three criteria which a group test must satisfy in order to qualify for school classification:

1. It shall have diagnostic value over a wide range of intelligence.
2. It shall be simple and easy to give and score.
3. It shall have reliable norms made on a large number of cases" (51: 55).

Trabue advises the use of the most reliable and most valid tests. If more than one test is to be used, then tests should be comparable (102: 425).

Granting that the Binet Test is the truest single means of estimate of intelligence, then "the correlation of any mass test with the Binet Test becomes of immense importance in estimating the value of the former" (76: 286).

Publishers of tests furnish lists of tests which conform to Dickson's essentials, and which meet the three criteria of Irwin and Marks. Such studies as Orleans on The Nature of Difficulty (71), Bailor on Content and Form in Tests of Intelligence (6), Cunningham on The Prognostic Value of a Primary Group Test (30), Goodenough on Measurement of Intelligence by Drawings (46), Jones on Effect of Age and Experience on Tests of Intelligence (53), and Vincent on A Study of Intelligence Test Elements (103), imply that newer tests will come nearer to meeting acceptable criteria.

An annotated list of group intelligence tests is published by the National Society for the Study of Education (108: 93-113). Torgerson lists tests published in 1928 (101: 376-381). Selection from the lists may well be made at present the basis of correlation with Binet results.

In the typical situation, tests were selected first for kimdergartens and grades one, two, and three. Age grade data of former years had shown that grade two presented a large retardation problem, and there was proba-

bility that organized testing in earlier grades and at least one grade following the second, might throw light on the retardation situation. The tests chosen were:

For kindergartens,	Detroit Kindergarten, and Pintner-Cunningham Primary Intelligence, Form A.
For Grades one,	Detroit Kindergarten, Detroit First Grade, Pintner-Cunningham Primary Intelligence, Form A.
For Grades two,	Detroit First Grade, Pintner-Cunningham Primary Intelligence, Form A, Haggerty Intelligence, Delta I.
For Grades three,	Haggerty Intelligence, Delta I.

Since group tests were to be used and since the testing was to be administered by newly trained teachers, a record of two or three tests in each grade was considered necessary. The tests chosen gave the kindergarten pupils the advantage of having one individual and one group test. Only one test was used in the initial survey in Grade III, because Grade III would later be given National Intelligence Tests.

Dougherty has made a Comparative Study of Nine Group Tests of Intelligence for Primary Grades, in the schools of Baltimore, Maryland. By means of a system of ranking of correlations of the nine tests with the Binet tests, she shows that certain tests give greater satisfaction in the matter of ranking the children by a group test than others, with the Binet as standard.

"For Kindergarten: Pintner-Cunningham is best.
For First Grade: Detroit comes first, then Pintner-Cunningham.
For Second Grade: Pintner-Cunningham.
For Third Grade: Haggerty, Delta I." (34: 106-107)

Table 3 shows Dougherty's correlations of the chosen tests, with Binet Tests.

Table 3

Dougherty's Correlations with Binet Test Results

For Kindergarten:

First Criterion 50 Pintner-Cunningham,
r = .803 with P.E. = .034 and I - K = .404
Second Criterion 186 Pintner-Cunningham,
r = .944 with P.E. = .006 and I - K = .670 (34: 33)

For Grade One:

First Criterion 50 Detroit,
r = .792 with P.E. = .036 and I - K = .390
First Criterion 50 Pintner-Cunningham,
r = .720 with P.E. = .046 and I - K = .306
Second Criterion 186 Pintner-Cunningham,
r = .887 with P.E. = .011 and I - K = .538
Second Criterion 186 Detroit,
r = .822 with P.E. = .016 and I - K = .431 (34: 39)

For Second Grade:

First Criterion 50 Pintner-Cunningham,
r = .607 with P.E. = .060 and I - K = .205
Second Criterion 235 Pintner-Cunningham,
r = .810 with P.E. = .015 and I - K = .414 (34: 50)

For Third Grade:

First Criterion 50 Haggerty,
r = .568 with P.E. = .065 and I - K = .177
Second Criterion 204 Haggerty,
r = .627 with P.E. = .029 and I - K = .221 (34: 59)

This table is read: According to the first criterion, tests of 50 children, the coëfficient of correlation of the Pintner-Cunningham Primary Mental Test with the Binet Tests is .803 with probable error of .034. The measure I - K indicates the improvement in estimate by a test over estimate by chance, which in this test is .404.

Pintner and Cunningham record coëfficients of correlation of their Primary Mental Test with Stanford Revision of Binet Tests as .71 in kindergarten with 20 cases; and as .82 in grade one with 19 cases. "The probable error of a score has been found to be 2 points. This shows the test to be very reliable" (72: 3).

Pintner shows a coëfficient of correlation between Detroit and Pintner-Cunningham scores of .67. This was derived from tests of 108 pupils (74: 55).

Stenquist has used Pintner-Cunningham and Detroit First Grade Tests for large numbers in Baltimore for several semesters (81: 79).

Haggerty has submitted his Delta I Examination to statistical treatment for the determination of reliability. The correlation between the standing in the tests given six weeks apart to 100 children of Grades one and two was .787 (Pearson)." (48: 54). In an unpublished letter to the organizer Haggerty states: "Different groups yield coëfficients for Delta I Examination from .68 to .78 with Binet."

On the basis of correlation figures the tests planned for the first three grades in Belmont were highly reliable.

The choice of a single routine test to be given late in the first year for grades from three through six, which should be reliable and thorough-going, lay between Haggerty Intelligence Examination, Delta II, and the National Intelligence Tests, Scale A, Form 1. Cattell finds a correlation coëfficient of .85 for the Haggerty Intelligence Examination, Delta II, with the Stanford Revision of the Binet (25: 441). Haggerty's figures for the same relationship are .84 ± .18 (47: 275). Root gives the coëfficient of correlation for National Intelligence Tests, Scale A, Form 1, with the Binet as .86 to .72 (76: 289). Two factors caused the choice of the National Test: first, the fact that so many of the pupils of Grade three ranked above and below the norms for the Haggerty Intelligence Examination, Delta I, made for dissatisfaction with a Haggerty Test among teachers who desired "more definite measure"; second, the National Intelligence Test is given and scored more easily (51: 56).

In summary, tests should be chosen which meet the known criteria of good tests, particularly in the matter of having been standardized through use with large numbers, and in the matter of reliability. Choice between equally valuable tests may be influenced by some local condition. Thorndike sees great possibilities of improvement

in tests and testing when "school authorities show their appreciation by taking the trouble to find out which tests are the best and to use them" (97: 349).

How can the teacher be brought into the plan?

It is possible for certain teachers who have had careful practice and training under guidance to administer intelligence tests. Some directors of tests go so far as to imply that unless a teacher can do such testing and interpret results in terms of needed teaching, she is not the right teacher for the times. (IX: Q4). In other words, measurement is an integral part of the new education, and a teacher must needs understand its administration and its import in order to rightly teach. Haggerty considers that "teachers will never have an adequate appreciation of the significance of intelligence or of the part it plays in the education of children until they learn to give intelligence examinations and to interpret the results of the tests. They should learn, through actual experience in giving and scoring tests, how great the differences of intelligence among children of the same age really are. Only when they deal with concrete cases measured in definite terms will they be able to direct their school work so as to secure the best results" (48: 8). "We have come to the point in the development of methods of testing intelligence where a good teacher can safely administer the tests. The average teacher cannot be expected

to make adequate psychological diagnosis of complicated cases", but "an intelligent teacher can be quickly taught to give and score the kind of tests described". (Haggerty Intelligence Examination, Deltas I and II.) (48: 8-9.)

Testing by means of standardized tests removes the good teacher temporarily from her habitual routine teaching, and permits her to set the situation and watch reactions of the child entirely according to his own ability. Right testing puts the child first, and places the teacher in a position to observe. The teacher's attitude, during the testing at least, must be scientific. Upon the teacher's ability to make the change from leader and helper to observer depends much of the success of her teaching.

Teacher administration of tests also does away with a former notion of measurement as uncomfortable pressure from above upon the teacher, pressure which was thought to result in little of value. Close coöperation in a test program makes the teacher a share-holder, and a sharer in the dividends from the program.

The involving of many teachers in the test program tends to strengthen coöperation, for it places all at work on a task which promises to be worth while. According to Maller, it is better to thus foster coöperation rather than competition (63: 163).

Clerical work, involving scoring, re-checking, and summarizing of data for their own class tests can be done by teachers, with no greater burden than is attendant upon the reading and grading of essay type examinations. The copying of final test data for office records can be done by teachers. If this work can be done by trained, clerical service, it is probably most satisfactory. In cities supporting research bureaus, skilled clerical service is available. The smaller city or town is obliged to train and utilize teacher service until the need for extra clerical help becomes apparent.

The experience of some directors leads them to believe that teachers should not administer intelligence tests, that these tests should be given by trained, experienced examiners only (IX: Q4). There are directors who, with reason, decry inaccuracies of teacher clerical work. Need for supplementary data on the matter of teacher administration of intelligence tests, and of teacher clerical service is evident, but the evidence of this study is that teachers can be trained to coöperate to the advantage of all concerned.

For two years, in the typical situation, teachers administered group intelligence tests, after they had been trained and guided by the organizer. Teachers scored the tests for their own, or neighboring classes, summarized the class data, and recorded much of the data upon office

forms. The table of Organizer's Time and Labor (Table 10), however, shows seemingly excessive hours during the first year spent in re-check and correction of teacher errors. As the re-checks and corrections were carefully reviewed with individual teachers and with groups, it is reasonable to infer that much of the time so spent was invested in possibilities of better work in the future. It was a necessary part of a program leading to coöperation and understanding by teachers.

The teacher's service in the program has been that of tester and recorder. As she has worked she has made discoveries of individual differences which make for better understanding of the child as well as of the facts behind classification.

CHAPTER IV

HOW RAPIDLY AND HOW CAN TEACHERS BE TRAINED TO ADMINISTER GROUP INTELLIGENCE TESTS?

Osburn, in answer to the question: "Ought teachers to be entrusted to administer intelligence tests?" says "Practically, some teachers ought never to be entrusted to give intelligence tests. Theoretically, any real teacher should be able to do it". (IX: Q4) This implies disparity between theory and practice. Whether the disparity is as great as it seems, is as yet an open question. That "the classroom teacher should test, provided the work is not too burdensome" is the decision of Stenquist. (81: 86) Right training, and patient, careful guidance will do much towards creating the interest, the understanding, the desire to do the work correctly, the observer-attitude instead of the instructor-attitude,-- all of which are necessary to successful testing. (3,4,5) Any real teacher should be able to do the testing under careful guidance. Terman emphasizes the need for teacher administration of both educational and intelligence tests, but he stresses the necessity for supervision by one of long experience in testing and in the interpretation of test data. (86: 291) Kohs reports a study made by Goddard at the Vineland laboratory, in which very careful

work was done in checking the Binet Test-learning of 58 inexperienced teachers. Ninety per cent of the work of the student testers was sufficiently accurate for all practical purposes. (57) The success of the student testers was due in no small measure to the painstaking guidance of Goddard.

The rapidity with which teachers can be trained to administer a given group intelligence test depends upon several factors: (1) Previous knowledge and experience of the teacher with standard tests; (2) Attitude of the teacher towards the test work; (3) Ability of the teacher to assume, temporarily at least, scientific attitude; (4) Available time for leader and teachers to meet for study together; (Size of system and organization of measurement personnel has important bearing upon this point. The small system has the advantage of opportunity for closer contact between director and teacher-tester.) (5) Ability of the leader to foresee the difficulties and possibilities of error in the given test, and to train the teachers to avoid these difficulties and errors.

The Belmont teacher-testers had had some previous experience with testing, though results had been unsatisfactory. They attended a general meeting at which objectives were set up. They were eager to learn and to do the work well. They were requested to study the manual

and the test before coming to the special training meeting. This special meeting of an hour and a half in length was given by organizer and teachers concerned, to intensive study of the one test in prospect. The special meeting was purposely set for the afternoon just before the test was to be given. At the meeting, the organizer gave the test to the teachers, that the teachers might get the "feel" of taking it. The teachers scored their papers, and summarized the data, as the summary sheets required it, all under guidance. Attention was directed to difficult parts of the test, to the necessity for using only the words of the manual, to the importance of accuracy in timing, to the requirement that the tester "keep away" from the child, giving him no aid whatever by look or act. Teachers seized upon opportunity for discussion. Each teacher "tried the test on some child that evening, and brought her questions from the trial to the principal the next morning before school. There was advantage in having the training period for each test, thus brief and intensive, and on the evening preceding the testing. Directions were remembered, and the spirit of the meeting rather than the spirit of former testing carried over into the next day's testing. Every test blank was re-checked and returned to the teacher for study. Criticisms and any evident errors were marked in red. The training for each test was conducted in a similar way at a single special meeting.

The training which the Baltimore teacher-testers experience, consists of: (1) Observing the test given by a trained mental examiner; (2) Discussing and reviewing the procedure; (3) Actual trial of giving the test under trained observation; and (4) Critical review. Then tests are given and re-checked, and X, Y, Z classifications made from data thus obtained. This is probably typical procedure. It is common observation (81: 79) that the interplay of the personalities concerned with the test program, has much to do with attitudes which make for success or failure.

What are the problem outcomes of the first testing?

To the organizer falls the responsibility of massing the data and making interpretations from them in the light of added contributions from teacher and principal. Teagarden shows that there must be critical evaluation of the test ratings; and that interpretation, in order to be of real value, must be made in the light of common sense knowledge of teacher and principal. (84) If the organizer is trained and experienced as he must be, his immediate problems after interpretation are concerned with how best to quickly set the found facts at work in the school system where they will give the best service. Problems a, b and c arose in the Belmont situation.

Problem a. How can the data be returned to the teacher, in order to be of immediate service?

It is possible to graph returns for schools, chart them, and present them at teachers' meetings; they can be mimeographed and sent to the principals or teachers. The best use of teachers' time, as far as many meetings are concerned, is worthy of serious consideration. Whether principals may be entrusted to make returns to the teachers, depends upon the previous experience of principals with test work, and upon the administrative organization. Whether the individual teacher, who ordinarily has little practical experience with graphs and statistics, will get from a presentation of data to a group of teachers as much as she could get from some simple, more easily understood presentation of the data immediately concerning her and her pupils, is the question. Whether a busy teacher will give mimeographed material the proper share of her attention, is also a question. It might be ideal if the organizer could carry to each teacher personally the returns from her work, at a time when question and discussion are possible. ordinarily time and size of system would not permit the use of this method.

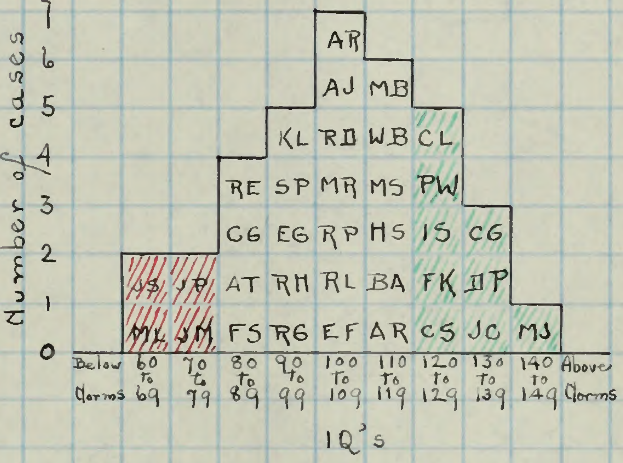
In the Belmont study a simple summary was devised, whose purpose was to center teacher attention upon the relative position of each pupil in her class, as shown by the particular test under consideration; and also upon the groups which appeared to be present in her class.

The bar graph yielded itself best to the need for a form which could bear the initials of each pupil. Bars containing the initials of pupils in the high group were colored green, in the low group, red, and in the central group were left uncolored. (See Charts I, II and III) This form and coloring were used for the initial report to the teacher from each test. With the report went a letter to the teacher asking her to keep the graphs within easy reach, and to note from time to time each child's position within his group. When several tests were given to the same pupils, the teacher was asked to note similarity of results for each individual child and to compare these test results with her own judgment of the child's ability as shown by daily contacts. The teacher was asked to keep the child's location in mind especially if he offered problems to which this matter of intelligence quotient might be a contributing element. In other words, the teacher was asked to accept the findings of the tests not as infallible labels headed "IQ" for indiscriminate use, but as one more set of contributory facts which might help upon occasion towards a better understanding of an individual child in a problem situation. Charts I, II, III show typical returns to a teacher of Grade two from three group intelligence tests. The making of the graphs for each teacher for each test was time-consuming. In terms of teacher understanding which came from the use of the graphs, it was time well invested. Teachers began to

Chart I

School __________ Grade ___ Teacher ______

Distribution of IQ's
Pintner - Cunningham Primary Mental Test
given
December, 1926

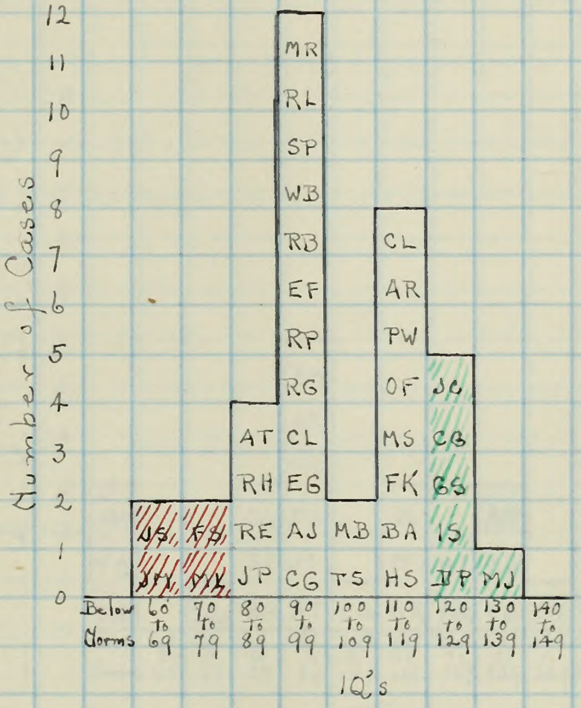


Range,	65 to 141
Q_3,	120.5
Med.,	106.4
Q_1,	91.5

Chart II

School .... Grade .... Teacher ....

Distribution of IQ's
Detroit First Grade Test, Form A
given
December, 1926

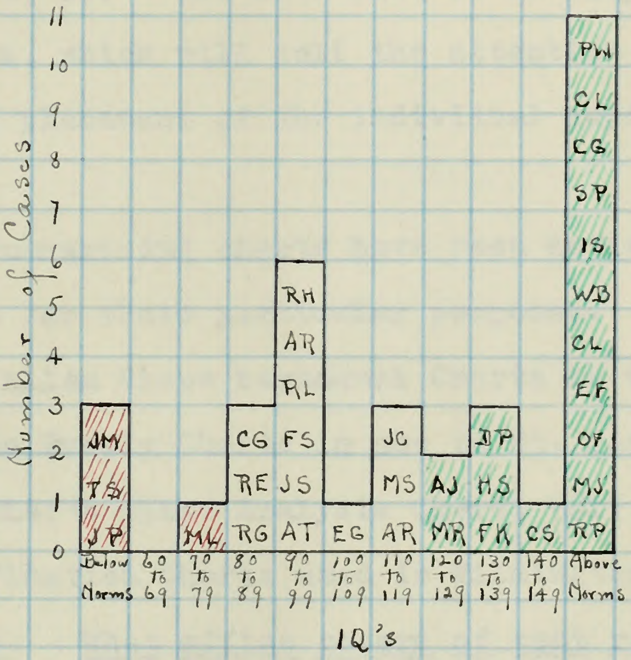


Range,		69 to 132
Q_3,	116.3	
Med.,	98.3	
Q_1,	90.8	

Chart III

School ____________ Grade ___ Teacher ___________

Distribution of IQ's
Haggerty Intelligence Examination, Delta 1, Form A
given
December, 1926



Range, Below to Above Norms	
Range, excl. of extremes,	71 to 146
Q_3, „ . „ „ ,	113.3
Med., „ . „ „ ,	100.0
Q_1,	83.3

discuss a child's ability in terms of test results as compared with teacher judgment or with some behavior of the child. The child and his right placing and adjustment became the center of thought. Teachers very early questioned whether curricular material was right for the pupils of their different groups.

An answer for Problem a, as derived from the experience in Belmont, is: The data must be in simple, easily understood form, which will call the attention of each teacher to the placement of the individual pupils of her room.

Various summarizing charts have been devised and found valuable for their particular purposes. Among these are the Allen Class Personnel Charts in use in Providence, the Everly Charts in use in St. Paul, the Baltimore Teacher's Class Analysis Chart, (81) and the Witham Classification Charts used in Wilmington, Delaware.

Problem b. What office record of test results is needed?

Each school system has developed its plan of office record of pupil data. In the larger cities which support research bureaus, very systematically planned and well-worked out systems of filing, take care of the data. In smaller cities with a lesser organization, there is need of centralization of school data on the individual pupil.

There are two methods of centralization in practice; one calls for separate files for health records, for subject report cards, for conduct records, for tests results, etc., all of which are available in the principal's office; the other calls for cumulative file of all of the data concerning the pupil in one place, and that place is at the school where he is a pupil. There is need for consideration of both methods as to practicability in a given situation.

The Baltimore Packet provides a single unit which follows the pupil from school to school. It contains, in card form, reports of school achievements, of tests, of guidance material, physical records, psychiatric report, etc. Stenquist demonstrated the packet in an illustrated address given at Schoolmen's Week, University of Pennsylvania, March 22, 1928. (83) The Yeoman's Envelope is in use in Revere, Massachusetts. The envelope contains all card and test data for the individual pupil and bears printed headings for a comprehensive summary on the outside of the envelope.

In the Belmont study each principal maintains a file for physical records, one for academic and attendance records, another for test records, etc. This arrangement is based upon necessary use made of the cards by school physician, teacher, director of tests, and others. Each person needs thus to handle only the cards of the depart-

ment with which he is concerned. There is evidence that ease of finding data, (the department method), is more than counterbalanced by good to the child which comes from each worker having immediate access to the entire case story (the individual method). There is not ample provision as yet for centralizing the complete records of the individual child.

The cumulative record of test results calls for a card upon which is recorded name and birth date of pupil, school and grade, test taken, date of test, chronological age when test was taken, score on the test, median score made by his class, standard score for his group, mental age, IQ, scholarship (a teacher rating). Space for other data which may be needed later. This is far from ideal, but has proved a valuable first form. Such form is needed as will preserve a summary record of successive tests in a way that makes observation of pupil progress possible. The actual test booklets contain much valuable information which should be filed within easy access and should be used by the teacher. The responses to the separate questions of a test call the teacher's attention to the individual differences of her pupils, while the summary data may give only relative positions of her pupils for classificatory purposes. The item argues for teacher testing and interpretation under guidance.

Problem c. Given the data from a test program, how can classifications be made? What is the next step?

This is the point at which much former work with tests stopped. It is the point at which those who are interpreting the data must realize the dangers of mechanization, and must keep in mind that no individual child can be wisely judged from the figures which he produces upon any one test alone. Lundberg reports the first attempt at classification in Flint, Michigan, as based upon three factors: intelligence, educational age, and teacher's marks on general school achievement. At present he is inclined to "disregard teachers' ratings, using a composite percentile rank from several tests". (62: 1) Any scheme of classification must make due allowance for elasticity of placement of individuals in the light of each bit of added data. This is also the point at which teachers need time to work with their pupils, comparing teacher judgment with test results. Much understanding comes from such comparison if time is given to think and work it over; much undue criticism of test results can develop if teachers are forced to accept test results as better than their judgments, and are hurried on to the next chapter of the test study.

Dickson at Berkeley and Oakland found that "no rule could regulate the classification of a school according

to the capacity of pupils to learn; no two schools can be treated exactly alike; the machinery must be subject to constant adjustment". He says that "to attempt to define what shall be done has a tendency to make for mechanical treatment of children, while what must be kept uppermost is that the individual needs of each child should be met as nearly as possible". (94: 35) If ability to do the work is the real criterion for classification, then mental age, as shown by intelligence tests, should give sound basis for placement. (64: 21-22) The general plan for placement was: "mental age to determine grade placement, intelligence quotient to determine the group within the grade", both of which were interpreted in the light of "health and scholarship data available". No general reclassification was attempted the first year though many individual pupils were adjusted in accordance with the findings, upon teacher request. The tests referred to in Chapter II were given early in the year to kindergartens and the first three grades. National Intelligence Tests, Scale A, Form 1, were given to grades from three through six in May, in order that grade placement in September, 1927, might be made upon the basis of objective measures combined with teacher judgment of scholarship. Test results and teacher judgment agreed surprisingly. Eight pupils, outstandingly wrongly placed, were adjusted that first year. They were pupils whom the tests found

capable of doing a higher grade of work; and, though teacher judgment opposed the advance placing, they have maintained their high placement with better conduct and scholarship ratings.

Pupils may enter the kindergartens of Belmont at the chronological age of four years and five months. This makes the median age for Grade seven months younger than is the acceptable age in most American cities, though Belmont's neighbor Boston does admit pupils to first Grade at five years of age. The problem of right criteria for school entrance is still only partially solved. Washburne uses the average of chronological age and mental age, admitting if this average is five. He thinks that the social criterion makes mental age alone inadvisable. (104)

First, grade placements were listed in readiness for September entrance, on the basis of mental age. Each pupil's test card was studied, during the summer, for proper grouping, A, B, or C, within the grade. This involved IQ and teachers' marks in conduct and scholarship. Table 4 shows the test data of one case each of low, middle, and high cases:

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

Table 4

Test Data of three Typical Cases

Case	Test	Date	Grade	Chron. Age	Score	Med. Score	Standard for Group	MA	IQ	Conduct	Scholarship
I Girl	Binet-Simon	1925	Kg.A	5-11				5- 6	93	D	C
	Detroit First Grade	11/'26	II B	7- 5	37	40	53	6- 8	90	C-	B-
	Pintner-Cunningham P.M.	12/'26	II B	7- 6	31	33	41	7- 1	94	C-	B-
	Haggerty A 1	12/'26	II B	7- 6	46	37	60.5	6- 8	89	C-	B-
II Boy	Binet-Simon	1926	Kg.B	5- 3				5- 2	98	B	B
	Detroit Kindergarten	11/'26	I B	5-10	19	15	25	5- 7	96	B	B
	Detroit First Grade	12/'26	I B	5-11	28	26	41	5-11	100	B	B
	Pintner-Cunningham P.M.	12/'26	I B	5-11	21.5	21	31	6- 1	102	B	B
III Girl	Detroit First Grade	11/'26	II	6- 8	57	49	53	8- 3	124	A	A+
	Pintner-Cunningham P.M.	12/'26	II	6- 9	54	40	48	9- 0	133	A	A∓
	Haggerty A 1	12/'26	II	6- 9	88	60.5	60.5	10+	above norms		-

It will be noted that two of these cases had been given Binet-Simon test previous to this study. This was the work of the State Clinic, so was retained as valuable data. In general, cases whose IQ's were near the 100 mark varied little from test to test. At the extremes of the distribution there was more variation.

Summary distributions from each test showed a rise in the curve at and above the 120 IQ point, and another rise at the 80 to 90 IQ point. The high were pupils from capable parents and favored environment. (91: 12) The 80 to 90 group were largely of Italian parentage. This suggested to the teachers questions as to how much weight language bore in the testing. Mead shows that group tests do not yield fair ratings for such foreign children. (66: 468). A group of Summer Session students at Boston University assisted in the study of the individual data. Otherwise it could not have been so thoroughly reviewed. Their notes on the three cases cited in Table 4 were as follows:

Case I is apparently functioning below the middle of her group. Haggerty A 1 test is probably not as satisfactory in low grade two as it is in a higher grade. The conduct difficulty should be investigated. Is basis for this to be found in the health record? She is slightly old for the grade. If correction of this conduct can be made, she is probably well placed in a B or low B section. Perhaps right placement may correct the conduct difficulty.

Case II is at just about the mid-point of the B group. The fact that the Pintner-Cunningham test involves language less than the other tests taken, may account for the slightly higher mark in that test. Properly placed in a B group.

Case III is slightly young for the grade, and scores high on all tests. Conduct mark is not quite up to scholarship. Is she finding work a little too easy? Why not try her in Grade three?

Cards of cases about which there was question were kept in separate file, and the placement made after consultation with principal, school physician, school nurse, or teacher. All A and C group cases were thoroughly studied. Eleven cases were noted as needing further clinic diagnosis.

When the classifications are actually made a certain amount of parental questioning develops. Belmont parents were very reasonable when they understood a little of the meaning of better classification. The contact of parents with the schools is closer in the smaller community. Baker of Detroit finds this no problem at all, as "measurement is counted an internal device of the school system". (IX: Qb) Size of city probably has much to do with the distance between parents and the school administration.

Classifications may safely be made on the basis of test data as modified by other available information. No classification can be permanent. As the child changes or grows, or as any new factors enter his school life, the

question of right classification again becomes paramount.

Teachers can be trained rather quickly to administer group intelligence tests if they are fairly intelligent, and are guided to an open-minded attitude towards the testing and its resulting problems. Their interest in individual pupils during and after the making of classifications is guarantee against mechanization of the classifications. Their attitudes and interests in individuals creates the necessary elasticity of classifications, provided the administration has the same desirable attitudes and interests.

CHAPTER V

HOW RAPIDLY CAN ALL TEACHERS AND PUPILS OF A SCHOOL SYSTEM BE INCLUDED IN THE PROGRAM OF TESTING AND CLASSIFICATION?

The question when put to the test in practice, resolves itself into: *How rapidly can all teachers of a school system accept the newer aims of education and come to realize values to be obtained from objective measures, enough to wish to share in their use and in their results? And how quickly can the necessary coöperations be brought about so that all pupils may be included in the program?*

The advice of Allen and Stenquist, written in unpublished letters to the organizer, is that an intelligence or educational test program must grow along with the felt need for it. If teachers of a particular building still cling to the older objectives of education and do not see values in the program, and do not wish to share in it, or do show inclination to oppose the movement, that school or building may be omitted from the program. Initial work may begin in a small way among those who at least wish to find out its values, perhaps in one school only. The testing work, as started in Belmont, has brought even the skeptical to the point of wishing to try it out.

The administrators of Junior and Senior High Schools were somewhat conservative in attitude towards a test

program when the program was instituted. Back of this conservatism lies the natural difficulty of the high school in shifting from the old historically founded, subject-centered view point to the new pupil-centered ideal. Accordingly, the initial survey was carried on in grades below the Junior High School. The giving of the National Intelligence Tests, Scale A, Form 1, to pupils of grades six in May of 1927 gave test ratings to pupils who were to enter Junior High School in September. The new administration in Junior High School grouped the entering class in 1927 on the basis of the test data without making explanation of the grouping to the teachers. Often through the school year teachers of grades seven remarked the evenness of their classes in ability, and the greater ease than formerly in the teaching of the group. In May, 1928, the giving of the National Intelligence Tests, Scale A, Form 2, in all grades six furnished data for classification of a second entering group in Junior High School. The same"evenness of ability, and ease of teaching" followed. By the end of the second year's work with the more homogeneous groups, the teachers were won to the side of ability grouping, and asked to know more of the tests and testing.

In the mean time the master of the Senior High School, through a special study, became interested in Terman and

Dearborn Intelligence Tests, and administered those tests in several classes where there were known to be pupils needing vocational adjustment. By the end of the third year, with the assistance of the new full time Director of Tests and Measurements, each pupil of Senior High School had taken an intelligence test and the data from the tests were recorded in the office file. It is probable that a fourth year will show added use of and interest in the data on the part of the Senior High School faculty. Table 5 shows a successful matching of ability and accomplishment which is to be expected of pupils of superior or very superior ability. The fact that these pupils can and are doing this successful work in the field of language, mathematics, and science agrees with Terman's findings in his studies of the gifted. (89: 285)

Ross' study shows significant correlation between composites of grade school factors and high school success. (77) In May, 1929, a tally was made of the actual repeating which had been done during the school life of the Senior High School pupils. Sixty-four of the 819 pupils had repeated one or more years, and had yearly records of good health. Scholarship or conduct, or both, were failing in High School. This had been recently recognized, and manual training, printing, domestic science, and sewing were added to the curriculum. Table 6 shows the wisdom of the move, in that the slow pupils succeeded in the less

academic program. It was enlightening to find upon the administration of intelligence tests, that the ability of all but three of the 64 fell between IQ's 73 and 97. Those three were between 100 and 115, and were social problems outside of school. The "80 to 90" group problem which the tests show in the early grades in Belmont is apparently still a problem in Senior High School years. Curriculum adaptations are bringing in a success element which is desirable. This is in keeping with O'Brien's study in which he found the highest per cent of failure among those studying mathematics, Latin, and German. (70) Data of this kind will attract a teacher's attention. If heads of research departments attempt to force the use of tests faster than the understanding and interest of the teachers is growing, discussion and opposition, rather than desire to find and utilize the facts is invited. The nature of the standard tests is such that pupils, after taking one test, look forward eagerly to the time when another may be taken. There is no difficulty in getting pupils to fulfill their part of the test program.

Table 5

Record of Ability, Health, Conduct and Scholarship
of Six Pupils
Whose Ability Tests Highest in 1928
in each of the four classes,
Belmont Senior High School

	Grades	Sex	Years Advanced	IQ	Health	Conduct	Sc	Hist.	Eng.	Lat.	Fr.	Math.	Ger.	Draw.
								Subject Ratings						
#1	IX	B	1	121	Good	A		A	A	B+	A-	A		
#2	X	G	1	147	Good	A		A-	A	B-		B		
#3	XI	B	1	131	Good	A			B			B-	A	A
#4	XII	B	1	157	Good	A	B+	A-	A-			A-		
#5	XII	B	1	145	Good	A-		B-	B	C		B		
#6	XII	G	1	141	Good	A		B	A-	B-		B+		

Table 6

Records of 18 High School Pupils who have repeated at least a grade

	Grades	Sex	Years Repeated During School Life	IQ	Health	Conduct	Sc.	Ehg.	Lat.	Fr.	Alg.	Hist.	Dom. Sc.	Draw.	Man. Tr.	Printing	Sew.	Notes
								Subject Ratings										
#1	IX	B	1	91	Good	D		D	D	D	D-							Left Sch.
#2	IX	B	1	73	Good	D		C+	D	D	D							Left Sch.
#3	IX	G	2	90	Good	D	C-	C-		C-	D+							
#4	IX	G	1	95	Good	C		C	C-		D-							
#5	X	B	1	97	Good	B		C-		D-		C		B	B			
#6	X	G	2	88	Good	B-		C			D							
#7	X	G	2	86	Good	B+		D			C		B				B	
#8	X	B	1	86	Good	D		C	C-		D							
#9	X	B	2	88	Good	A		C	C-	D-					A+			
#10	X	B	3	90	Good	B	C+	C-			D					B		
#11	X	B	1	86	Good	B-		B		C-	C	D				B		
#12	X	G	1	86	Good	B		C			D	C-	B				B	
#13	XI	B	2	81	Good	A		C	D	C-		D			B			
#14	XI	B	1	91	Good	B		C-		C-	B+	C-				B-		
#15	XI	G	3	81	Good	B+		C			C-		A-				B-	
#16	XI	B	2	87	Good	B		B		C-		C-				B		
#17	XI	B	2	82	Good	C		B+	C	C-								
#18	XII	B	1	93	Good	B		D-		D		C-			B			

Complete data was not available for the other 43 repeaters. The incomplete data imply that the 18 cases tabulated are typical. Success in the more practical subjects followed the facing of failure in the more scholastic subjects.

What problems follow the first classification?

Initial problems after classification have to do with teacher and pupil adjustments to the new situation. In the small school system parents are interested to know a little of the whys.

There are still teachers who clamor for opportunity to teach a superior group, not recognizing that the inferior group is as great a challenge to their ability. See reason suggested by Kallom. (IX: Q8b) The administration can guide the desire to accept the challenge. Teachers at work in the classified groups feel immediately the need for curricular adjustments, and they think and work much along those lines. Pupils adjust themselves quickly to the classified group, and soon utilize time which was spent in the unclassified group in teasing inferiors, or in succumbing to the wiles of the superior, in matching their own strength with that of equals in ability. The struggle for superiority in the equalized group is remarked early by the teachers as a healthy development. (44: 26-28) New needs are uppermost in the minds of all the workers. Needs demand prompt recognition. Further work requires recognition of the most urgent need.

Where is the best place to begin intensive work, and what can be done?

According to Gesell's recent studies which show that tests of intelligence can be made very early in life,

there is little danger of measuring too early in school life. (38 and 39) The earlier individual differences are discovered and measured, the sooner adaptations may be a preventive measure. Dickson advises that "Every child should be given an individual test as soon after entering school as possible, not later than the middle of the first term if he speaks English". (32: 44)

In September, 1927, in Belmont, pupils below Grade eight had been placed as intelligence tests, plus health and teacher ratings of conduct and scholarship, plus other informations had indicated. The new entrants into kindergarten and first grade were untested, and teachers recognized differences which they wished measured. Of 337 pupils of Grade one, 41 or 11.9 per cent were repeating. Table 7 shows this to be the highest per cent of repeating in the grades below Junior High School.

Table 7

Promotion Data - June, 1927

Grade	No. Promoted	No. not Promoted	% Promoted	% Not Promoted
Kg.	292	4	98.4	1.6
I	296	41	88.1	11.9*
II	278	23	93.06	6.94
III	277	21	92.89	7.11
IV	259	6	97.5	2.5
V	263	10	96.25	3.75
VI	265	6	97.63	2.37

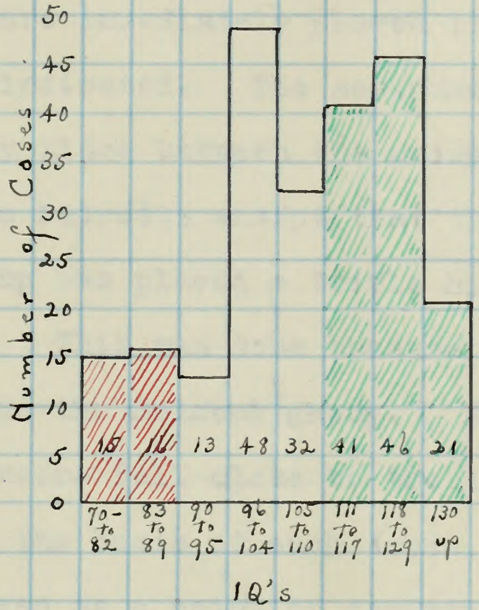
That this is not an unusual situation is shown by Reed who investigated practices in first grade admission and promotion. "Comprehension by administrators of the significance of first grade failure is evident, in that 192 out of 219 of them checked one or more procedures used in their schools or school systems to eliminate or reduce (first grade) failure." (75: 116) Reed in summing up present tendencies says: "The use of chronological age for Grade one admission shows tendency to hold to traditional objectives, disregard the findings of scientific research, and ignore the value of kindergarten learning as a basis for Grade one work on different levels of ability." (75: 121) "Not to use objective measures when available is failure to accept what science has to offer." (75: 122) Keyes has shown that repeating a grade does not result in any permanent improvement of scholarship of the arrest. (56: 63) Broady finds that "failure and repetition as a policy are not acceptable". "When the standards have been set beyond the pupils' abilities, the pupils should not be required to suffer." (19: 37) It was evident that the school entrance situation in Belmont as elsewhere should be studied.

The giving of Stanford Binet tests was a desirable move. Opportunity to learn to give this test, at least for the early years before nine, was offered at the opening of the second year to all teachers of kindergartens

and grades one. All of the 21 teachers voluntarily undertook to learn the tests, and to test their pupils. They studied the directions with the help of the organizer, and tried to test one child. The one record blank was thoroughly reviewed by organizer and teacher together. The second week, two pupils were tested; the third week, five, always under guidance. After ten successful tests were made, teachers were permitted to complete the tests for their classes, but each test blank was subject always to re-check and criticism. Tests in the early years of the Binet scale are comparatively simple, making this procedure satisfactory with far less teacher training than must be given if tests are to be made higher up in the scale. Chart IV shows the findings of the 232 tests in terms of IQ.

Chart IV

Distribution of IQ's
of the
232 Kindergarten Pupils of 6 Schools
as shown by
Stanford Revision of Binet Tests
Given Nov. 7, 1927 to Jan. 10, 1928



Groups,	Sp.	C	B-	B	B+	A-	A	A+
No. in group	31			93			108	
Belmont groups	C			B			A	
Detroit groups	z			y		x		

These 232 pupils were the entire group of kindergarten pupils, entered September, 1927.

These tests were administered by 21 different teachers.

The Binet questions used with Italian pupils were those which Terman used with children of foreign parentage. (89: 22-23) The results showed that language handicap was probably not responsible for low IQ's among certain of the Italian children. This fact is supported by Bere's study. (11)

Pupils were immediately placed in A, B, or C groups as the tests indicated. The assignment of the more or less arbitrary lines between the groups paralled the X,Y,Z plan in use in Detroit, except that the center of the Belmont A group was placed a trifle higher than the Detroit X. (7: 103) This was done because of the largernumber of high IQ's in the Belmont group. No child, whose intelligence test score fell close to the line between groups, was placed on the basis of such score alone. The test score was "taken as a point of departure for further study of the pupil". (94: 25) Attempt was made to avoid any arbitrary placement and to give the child the benefit of any doubt.

In summary, at school entrance is a fertile field for intensive first study. Objective measures, applied by teachers as soon as possible after school entrance, yield results in teacher understandings of retardates and accelerates. Classifications which help the child to work on a level with his ability as shown by the tests can then be made early, with less probability of massing of

retardation in Grade one, and also with no possibility of the child's presenting the same unrecognized problem in Senior High School.

From experience in Belmont, the answer to the question: "How rapidly can all teachers and pupils of a school system be included in the program of testing and classification?" would seem to be: "As rapidly as teachers can come to accept the newer aims of education and to understand and appreciate the values of objective measures of the child. The understanding comes slowly over a period of years. Stenquist has planned for an eight year program in Baltimore. Allen has been at work slowly and steadily for five years, and is far from completing the plan for ability grouping in the Providence schools. These men are capable workers. All may profit by their experience.

CHAPTER VI

DEVELOPING THE NEED FOR ASSISTANCE

How rapidly can consciousness of needs be developed?

It has not yet been possible to plan a test program which will function satisfactorily in every detail without the making of manifold adjustments. As the initial work progresses, needs develop all along the line. Adjustment of the details of minor needs are easily made. Major needs grow and press until their weight is felt. Consciousness of need is apt to lag behind the need and develop only when the need presses hard enough to threaten friction if it is not adjusted. The weight may be well on before realization comes.

Needs developed rapidly as the work progressed, especially during the second year. (100) These needs found their way into consciousness near the close of the year when National Intelligence Tests were given to grades three through six. Returns came in more slowly than at the close of the first year. There were questions as to whether the best use of teachers' time was being made in the scoring and re-checking of tests, and in recording the data. Should not these mechanical matters be attended to by clerical help? The organizer, or some trained worker needed more time for closer personal contacts with the

teachers and the schools. Questions needed to be answered before they developed into problems. Classifications in the making, needed more oversight than they were getting. This was especially true because of the rapid shift in school population. It was possible for children to enter the system, attend school a short time and move away again without being tested. (A check made in April, 1929, showed that of 1089 pupils who were given National Intelligence Tests in May, 1927, and of 912 who took the second form of the same test in the same grades in May, 1928, only 432 pupils remained who had taken both tests.) There was a need for some test of new entrants to determine right initial placement. Teachers needed to be relieved of certain pressures of detail; principals needed help in defining their problems; the organizer needed far more time for the administrative task, and could afford to turn over details to others. Those concerned with each need felt and expressed the urgency of it. Adjustment was necessary.

What are the possible and best plans for extra assistance?

There are at least two ways of arranging for the improvement of the situation as described in question 1. Both of these ways involve policy and therefore require conference with the superintendent. (2: 260) (1) Extra part-time assistance can be planned to care for detail

work, such as clerical matters; extra teaching force can be engaged in order to release time for teacher administration of tests according to the initial plan. (2) Full time skilled assistance can be engaged for direction of tests and measurements, and clerical force can be enlarged as it becomes necessary.

Plan 1 savors a trifle of pressure upon the teacher, and so has possibilities of undermining the coöperation of all, which is so important to the success of a program. Plan 2 brings a new skilled person into the coöperative scheme, one who can relieve the over-pressure upon the organizer and teachers and still have ample time and opportunity to make personal contacts in the system wherever it is needed for assistance or to avoid friction. Plan 2 was considered the more desirable for Belmont, and a skilled leader entered the service in September, 1928, as full-time director of tests and measurements.

How can the expense of extra help be justified?

Retardation is a tremendous expense to a school system, and shows no credit side to its ledger. (56: 63) On May 1, 1929, a checking was made, from school record cards of each individual pupil, of the actual years of repeating which had been done by the pupils then in the Belmont grade and high schools. Table 8 shows the resulting data for each school for each year, yielding a total

of 526 years of repeated schooling. This repeating was done by 395 pupils.

Table 8

Years of Repeating
Done by 395 Pupils Enrolled in Belmont Schools
May 1, 1929

Schools - Yr.	'16	'17	'18	'19	'20	'21	'22	'23	'24	'25	'26	'27	'28	Totals
Senior High School	2	5	9	29	17	11	18	17	2	5	12	7	2	136
Junior High School		1		2	18	21	23	23	24	25	16	4	0	157
Butler		Partial				2	4	6	10	8	17	6	8	61
Kendall		Area					3	6	8	17	15	14	10	73
Homer		of					7	4	11	11	8	2	3	46
Wellington		Shift in Popula-					1	1	1	3	5	0	0	11
Payson Park		tion					2	2	5	6	6	1	9	31
Chenery	—	—	—	—	—	—	—	2	1	1	5	0	2	11
	2	6	9	31	35	34	58	61	62	76	84	34	34	526

Table 9a shows the actual amounts of money spent by Belmont in this poor investment. The "No. of years repeated" is listed from Table 8. The calculation of cost is based upon the "per capita" cost for each year, as listed in Belmont School Reports for the given years.

Table 9a

Cost to Belmont of Actual Retardation

Year	Number of Years Repeated		Per Capita		Cost
1916	2	x	$35.72	=	$ 71.44
1917	6	x	37.68	=	226.08
1918	9	x	41.80	=	376.20
1919	31	x	45.95	=	1,424.45
1920	35	x	49.89	=	1,746.15
1921	34	x	70.57	=	2,399.38
1922	58	x	75.51	=	4,379.58
1923	61	x	79.19	=	4,830.59
1924	62	x	79.33	=	4,918.46
1925	76	x	82.86	=	6,297.36)
1926	84	x	85.33	=	7,167.72)
1927	34	x	87.19	=	2,964.46)
1928	34	x	91.76	=	3,119.84)
Totals	526				$39,921.71

This table is to be read: In the school year of 1916-1917 two years of schooling were repeated by pupils who are still in school in Belmont. At the per capita for 1916 of $35.72, Belmont paid $71.44 for the repetition.

Chart V shows the curve of the actual number of years repeated during each school year from 1916 through 1928-1929. The sharp drop at 1926 is to be noted. This was

the time of the initial work in classification according to test and other findings. So far as is known, there was no other unusual happening within the system which might have caused the drop.

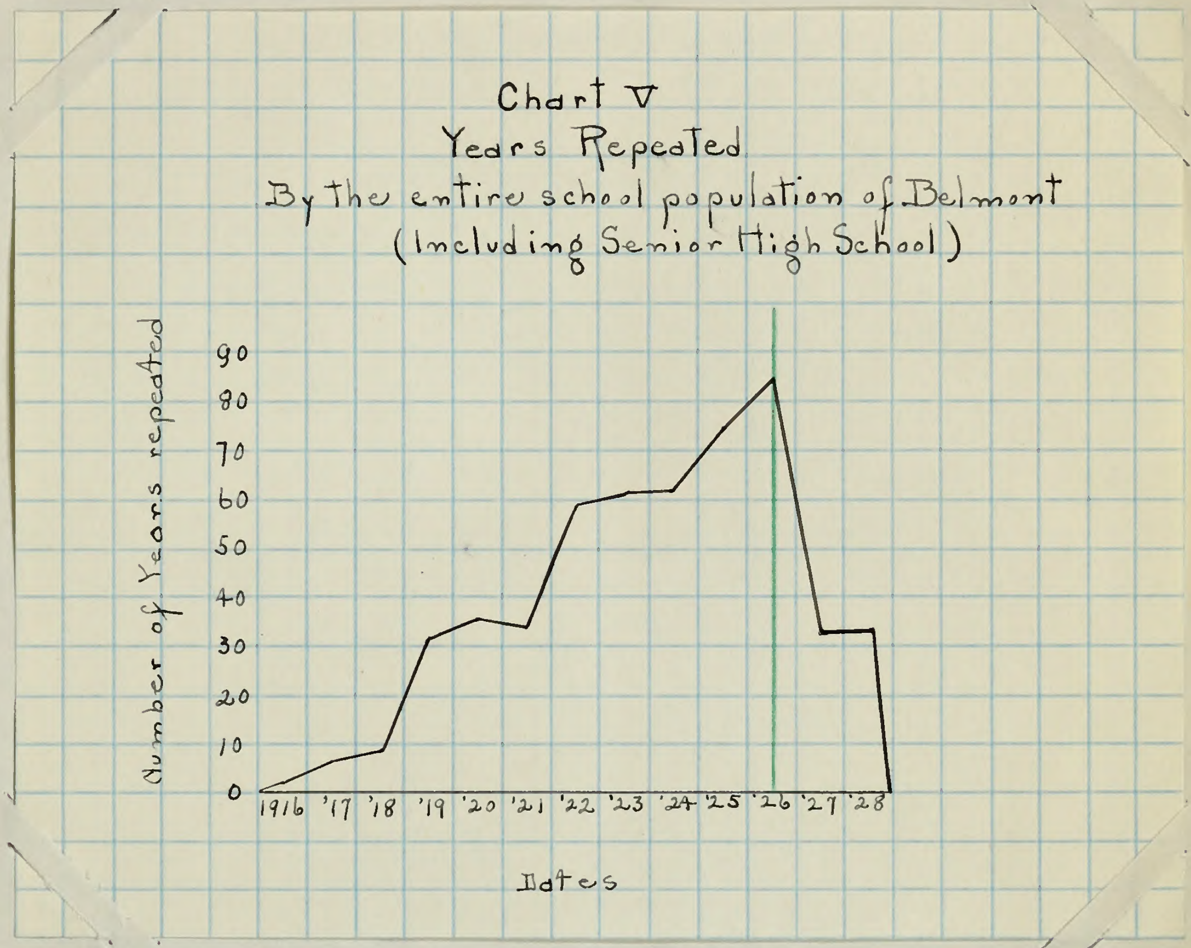


Chart V
Years Repeated
By the entire school population of Belmont
(Including Senior High School)

From Table 9a it appears that the cost of actual repeating for the two years just preceding the test program, 1925 and 1926, was $13,465.08; for the two years following the beginning of the program, 1927 and 1928, the cost of retardation was $6,084.30. This yields an actual saving of $7,380.78, a tidy salary for a director of research in a small city school system, as well as ample funds for materials for the test program. These

figures are based upon the simple years of marking time for the pupils. The loss to each pupil through realization of failure is often spread over future years. This loss is as yet unmeasured in dollars and cents. McCall counts a retarded pupil as $1000 lost to the city. (65: 4) Then Belmont's loss from repeating, as figured by McCall, is $395,000.

In addition to costs of retardation, must be considered amounts of money saved through acceleration. Table 9b shows actual savings through advancing 43 capable pupils.

Table 9b

Savings for Belmont from acceleration

Year	Number Accel. 1 yr.	Per Capita	Saving
1927	5	$87.19	$ 435.95
1928	30	91.76	2,752.80
1929	8	91.44	731.52
	43		$3,920.27

If the saving from retardation covered costs of test program and its direction only, then there would still be a gain of $3,920.27 from right placing of superior pupils.

What are the new problems of a new worker?

In order to meet problems of the test program presented by superintendent and organizer, as well as those

of principals and teachers, it was necessary for the new director to not only have but personify the service view point. The third year's work, except for the giving of Stanford Revision of Binet Tests to all newly entering pupils in kindergartens and grades one, was devoted by the director to giving service where it was needed. This service proved to be needed in five different lines:

(1) At the opening of the school year the director took over the administration of the Binet tests for entering pupils, thus relieving the teachers of some of their question as to time for doing it, and also making time for discussion, with the teachers concerned, of the problems which the individual pupils tested presented.

(2) Because of the unusually large shift in population the card file of test records was incomplete. It became her duty to check these records, to see that every pupil in the system took an intelligence test, and to record the results.

(3) As each new pupil entered any grade,he was tested, studied, and placed in school as his record indicated.

(4) Junior High School records were incomplete. The director made a thorough study, and many re-tests of the pupils, and arranged complete re-classification into ability groups.

Assistance was given in completing the intelligence records of pupils of the freshman and sophomore years of

Senior High School. Dearborn and Terman Group Tests were used.

(5) Throughout the system were problem cases, located by the tests or by teachers, which needed attention. This required that much time of the director be devoted to coöperation with the State and other clinics, and with medical and social agencies. While this work was going on, there was opportunity for service to the principals in making of classifications in the light of test findings. In two schools the use of Allen Classification Charts and Cards has proved of value, in graphically calling attention to wrongly classified pupils.

The first year's problems of the new worker have to do with getting acquainted with and fitting into the system, and with evening up the former rough and weak places in the measurement work and its records, so that the following year's work may run smoothly according to plan.

CHAPTER VII

HOW CAN THE ENTIRE ORGANIZATION BE INVOLVED IN WORK UPON THE PROBLEMS OF CLASSIFICATION OF PUPILS?

Can blunders be avoided?

Webster's dictionary defines "blundering" as "a stumbling, a floundering". It would seem then that the question is one of whether it is possible to carry a test program through to its ultimate application without making gross errors, or without reaching aimlessly for an elusive goal.

Forcing a situation is apt to invite debate and opposition. The laying out of objectives which can be easily understood, and the motivating of the tasks in such a way that the workers are interested to volunteer to do the work, will guarantee a necessary interest background. Guidance must be expert, and planning must be in the light of the best information available. There must be a willingness to accept or discard as data dictate. The administrators must be quick to sense the needs and to try to meet the needs promptly in the best known way. Early conference upon problem matters has been found in Belmont to be an effective means of bringing desired coöperations, as well as of clearing up difficult situations. Each person concerned in a test program and in its applications plays an important part in the whole scheme. Blunders

can be reduced to a minimum if leadership is fore-sighted and wise and if all followership is intelligent and loyally interested in working coöperatively towards well-defined goals. These conditions of success are not easily attained. In Belmont the entire first year of work was devoted to creating right soil in which these conditions might develop. Decreasing amounts of work were needed in the second and third years, as understanding and coöperation developed, and as response came for needs. Stenquist has found the major part of three years necessary for rightly "selling" the program to the teachers. (IX: Q3)

Can coöperation be secured?

Conference has been mentioned as a means of forestalling blunders. In the large school system which boasts a department of research, routine matters are mechanized, and policies and powers more definitely laid out and assigned, than in the smaller system. The administration and the child are farther apart in the large system. Many matters which were brought to conference in the small system might not arise in the large, or might be cared for there by a different method. The fact that the organizer was a research worker and not a paid employee of the school department made certain definition of power necessary also. During the three years of this study these matters were considered at conference of superintendent and organizer:

(a) Values from former tests, if any.

(b) Objectives for the test program.

(c) Planning of the initial test work, as to quantities of materials, costs, meetings needed, method of attacking the problem of slow return of data from teachers and principals.

(d) Study of graphed returns from tests, showing presence of two major problems: (1) the large number of superior children, and (2) the large number of Italian children at the 80 to 90 IQ point.

(e) How to get the clerical work done?

(f) Study of types of filing, and development of a test record card for temporary service.

(g) Times and method of making the actual classifications.

(h) Promotion policy.

(i) Under the conditions, what can best be done for the superior children?

(j) What more can be done for the Italian group who appear to be below normal according to the tests?

(k) The school entrance situation, and what should be the admission policy?

(l) Case of one problem pupil.

(m) One case of adjustment between teachers.

(n) Test policy for Junior and Senior High School.

(o) Is curriculum revision imminent?

(p) Need of a full-time director of tests.

(q) Matters to be placed in the Annual Report anent the measurement work.

(r) What are the most pressing questions about the measurement program? What answers can be found from those of more experience?

In September, 1928, a director of tests and measurement was appointed to the Belmont schools. Conferences of organizer and director were devoted to:

(a) Reviewing of the program of the first two years as to objectives, accomplishments, needs, weak places, difficult places.

(b) Passing on of superintendent's policies, and planning for any needed new policies.

(c) Each new step in the program thoughtfully worked over, and planned.

(d) Reports of work as it progressed.

The director attended all conferences of superintendent and organizer after her appointment to the service.

Conferences of organizer and principals were upon these matters:

(a) The routine matters of the testing.

(b) Details of classification.

(c) Adding of principal's data for making placement of problem cases, or cases about which tests did not closely agree.

(d) Teachers' errors in test work which principals could help to lessen.

These included:

- Uniform method of writing dates.
- Use of child's full name upon all reports (no nicknames.
- Figuring of ages.
- Accuracy in adding, and in other calculations.
- Care not to reverse the fraction $\frac{MA}{CA}$ in finding IQ.
- Handling of erasures.

(For dates of conferences see Table 10.)

At meetings of superintendent and principals opportunity was offered for questions as to policy or test program. These were usually on routine matters of assignment of particular duties connected with the test program.

Principals and teachers conferred on the details of administration and summarization of tests and results, and upon details of classification.

Director of Tests was available daily during the school year of 1928-'29 for conference with teachers upon any untested or unsatisfactorily classified pupil, problem cases, and on further study of, and need for tests.

Each person concerned worked with others most coöperatively in the study of the best placing and the best functioning of each individual child.

Can the entire organization be involved? If so how, and at what cost?

In two years the entire elementary school force in Belmont was involved, and the Junior High School was favorably inclined towards the intelligence testing movement as an aid in better classification. Certain tests had been given in Senior High School. At the end of the third year, the Junior High School was thoroughly classified, and teachers appreciated advantages of the grouping. At Senior High School accurate and valuable data from tests were on file, and were used in making curricular and vocational decisions, with promise of a growing teacher interest in homogeneous classification.

The method of involving each person in the system was that of interesting each and all in such a placement of each individual child that he might have opportunity to function at as high a level as his tested ability indicated. This was apparently accomplished by careful planning and direction and coöperation, all three of which were lacking in earlier unsuccessful attempts at testing, due to lack of provision for necessary help.

Refining of the classifications through use of the Allen charts has been more thoroughly done in two of the larger schools under the personal direction of the Director of Tests and Measurements. Extension of this thorough work promises to strengthen coöperation.

The cost in time and labor of the organizer appears in detail in Table 10. To these facts should be added a year of full time work of the director. Outside of a few minutes at regularly scheduled meetings, item 3 of Table 10 is a measure of the time given by the superintendent to the work. It appears from Table 10 that clerical help would have released time for the organizer to attend to other parts of the test program. The 47 hours given to checking group tests and noting errors for teachers (item 11), was far from being a waste, as appeared in the more careful work of teachers in later testing. The 42 hours spent in making graphs (item 16) was well invested, in that the graphs were objective means of centering teacher thought upon the best placing of the individual child.

Table 10

Hours of Labor and Teaching
Given by the Organizer
From September 1, 1926, to June 15, 1929

	1926				1927												1928											1929						
Labor	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June		Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June	Totals in Hours
1. Reviewing former test results		2.50														3.	1.																	6.50
2. Outlining objectives and testing program	6.	4.										4.								2.		2.		2.					3.					23.
3. Conferences with Superintendent		1.	2.	1.		1.	.50	1.	1.	1.			2.25							1.		1.		1.	2.	1.		3.50		1.50				21.75
4. Conferences with Director of Tests																								1.	6.		1.50	2.50	1.50	2.50			2.	17.
5. Conferences with Principals						1.	1.50		1.												1.50					1.	8.			6.				20.
6. Conferences with Teachers on Tests					1.	1.									15.									2.							2.	10.		31.
7. Special Conferences with Parents							3.	.50	.50					.25																				4.25
8. General Teachers' Meetings		1.50																																1.50
9. Group Teachers' Meetings			4.						1.						1.																			6.
10. Giving of Special Binet Tests							2.	1.																										3.
11. Corrections on Group Tests				2.	12.	10.	2.		8.	4.												6.		3.										47.
12. Corrections on Individual Tests															9.	5.	10.																	24.
13. Adjusting of Problems of Teachers				.50										.50																				1.
14. Work on data with Director of Tests																										4.	1.							5.
15. Making Charts and Graphs for Superintendent					2.	3.	.50			.50								5.			9.			1.										21.
16. Making Graphs for Teachers					15.	15.	2.			10.																								42.
17. Obtaining data at Superintendent's Office			.25	.50																								.50						1.25
18. Study of data for Classification										12.	13.	8.						3.	2.			2.						1.						41.
19. Making lists for use of Principals in Classification										10.								1.	.50															11.50
20. Summarizing data from Tests				6.	13.					4.								2.	1.					3.				3.						32.
21. Clerical work on Record Cards and Office data							10.	11.	10.	8.	10.																				20.75	22.75		92.50
22. Writing 157 letters																																		39.
Totals	6.	9.	6.25	10.	43.	31.	21.50	13.50	21.50	49.50	23.	12.	2.25	.75	25.	8.	11.	11.	3.50	3.	10.50	11.	39.	13.	8.	6.	10.50	10.50	4.50	10.	22.75	32.75	2.	491.25 or almost 82 six-hour days.

The original estimate for cost of test materials was $400 a year. Actual costs of the program are shown in Table 11.

Table 11

Cost, in Money expended, of Testing Program for 3 Years

Test Materials:	
First Year	$ 230.95
Second Year	107.45
Third Year	56.50
Director's Salary for Third Year	2,000.
Postage, cards and other incidentals	15.
Total	$2,409.90

Terman and Dearborn Group Tests were on hand from a former supply. Their cost is not included here, as they would otherwise have been wasted. From conservative figures, the drop in retardation saved Belmont $7,380.78; the right placing of accelerates saved $3,920.27. The total saving was $11,301.05. The cost in actual money expended was $2,409.90. Deducting cost from saving, the profit from classification in money was $8,891.15.

CHAPTER VIII

WHAT ARE THE COMPENSATIONS FOR THE EXPENSE OF TIME, LABOR AND MONEY?

According to Irwin and Marks, "Psychological classification does not encourage the idea that only the best intelligence is worth while and the rest nowhere; on the contrary, it leads us to value all kinds and grades of intelligence and to try to understand how to develop them towards a life of personal effectiveness and social usefulness." (51: 163) Goddard says that "the social efficiency of a group of human beings depends upon recognizing the mental limitations of each one, and upon so organizing society that each person has work to do that is within his mental capacity and that at the same time calls for all the ability that he possesses". (43: 127) Terman et al find that "Universal testing is fully warranted, whether considered from the point of view of money cost or labor cost. If we can afford to spend $50 to $100 a year for a child's instruction for ten or twelve years, we can surely afford to spend 6 to 10 cents for a test whose results may affect the child's entire educational career. The labor cost is too small to be counted an obstacle, now that satisfactory group tests are available. If the results are properly used, the cost is a negative quantity, for in the long run labor is saved". (94: 22)

From the work of Davis at Jackson, Michigan, it appears that under ability grouping "pupils are happier and more successful at their work". (31: 142) Worlton makes the same finding, (112: 273) as well as the following: More rapid progress in subjects, more congenial leadership, greater stimulus for individual effort, better chance to escape snobbishness and conceit, fewer "disciplinary" problems, better teacher morale. (112: 273-274)

Torgerson shows from his data at West Allis, Wisconsin, under ability grouping, (1) failures are reduced ("40.9 per cent reduction in first six grades after three years of classification"); (2) "proper grade placement and classification reduce maladjustments and mental and educational over-ageness and under-ageness, and result in a better educational product"; (3) "proper grade placement tends to raise accomplishment quotient to normal maximal efficiency". (100: 180)

Layton of Detroit was "gratified by the constant and substantial increase in the number of group mental tests in Detroit because it reflected a great interest on the part of the teachers and principals and because the teaching public shows an earnest desire to make use of the test results". (60: 130)

It has been the experience of the organizer that teachers who have worked upon data having to do with ability grouping are more prompt in response to a call for

data than teachers who have not had that opportunity to coöperate. The organizer had occasion to ask for some simple data from two schools in a suburb of Boston on November 23, 1929. No time was set for the return of the data. (For obvious reasons the names of schools and city are omitted.) One of the schools was homogeneously grouped, the other not so grouped. Teachers of the two schools were matched as to training and teaching experience. The principals are apparently well-matched as to executive ability. The grades and types of pupils are matched. Data from the grouped school were returned with no comment on December 5, 1929, after seven school days, though the Thanksgiving recess had interrupted work upon the data. The data from the ungrouped school were returned on February 9, 1930, after much explanation of the "very pressing, busy times". There was at least enough teacher interest in a matter of data in the grouped school to bring prompt response.

Maller presents data for the finding that "Members of a group will tend to show a higher degree of coöperation if there is, within the group, little variation in intelligence, age, and social factors. A maximum of coöperation is obtained when a group is of the highest homogenity. The sectioning of classrooms on the basis of intelligence and other factors would thus augment the cultivation of coöperativeness". (63: 163)

To summarize the findings of the workers quoted, credits include the following: With ability grouping, pupils are happier, make more rapid and successful progress in subjects, enjoy greater opportunity to develop leadership, find greater stimulus for individual effort, and have a better chance to escape snobbishness and conceit. The school has a more congenial social atmosphere and presents fewer disciplinary problems. There is an increase in teacher interest and morale. Pupil failure is reduced, over- and under-ageness is reduced and accomplishment quotients are raised. Better opportunity is offered for cultivation of coöperativeness.

The credit account in Belmont shows that fewer disciplinary problems occur after homogeneous grouping, especially in Junior High School. When the eight seventh grades were grouped, the highest group was found to contain the "rascals" of the grade, "as carefully segregated as though the testers had known and placed them because of their rascality". The principal suggested the advisability of re-grouping this class upon some other basis than that of test scores. He was advised to give the group a skilled teacher who knew the situation. This was done. The class settled down to business immediately, and led in extra-curricular activities throughout the year, with no evidence of disciplinary problems. Increase in accomplishment is shown by the fact that all of the 43

pupils advanced by the classification have succeeded in the new placement. The fourth year, planned for intensive program in educational measurements, promises much data on this point.

Chart V gives evidence of substitution of success for failure in the immediate drop of the curve of retardation at the time of classification. See page 73. Further evidence appeared when, in May, 1929, the data of the 432 pupils of grades three through six, who took National Intelligence Tests in 1927 and again in 1928, were studied. Of the 432, only 5 repeaters were found. 427, or 98.9%, were comfortably and successfully making the grades in which they were placed. The five cases investigated individually, proved to be cases whose difficulties had social origin.

An investigation of teacher interest was made in May, 1929, by means of the following questionnaire. Answers were returned very promptly by all of the teachers concerned.

(Note: Please use √ for "Yes")

School ....

Grade ....

Teacher's Name ....

1. Were you one who gave Binet Tests to pupils of Kg. or Gr. I in October, November and December, 1927? ...

2. Have you given any Binet Tests since February 1928? ....

If so, how many? ....

3. Have you called cases to the attention of Director of Research, State Clinic, or Social Service, since February, 1928, of your own initiative? ....

How many? ....

Would you have done this if you had not had experience with tests? ....

4. Have you made use of the information which is on the test record cards in the office file? ....

What use? ....

In how many cases? ....

5. Have you asked to have group intelligence tests given to your pupils since May, 1928? ....

6. Have you made use of any facts about pupils, which the testing program brought to your attention? .....
....

What facts? (Please use no names) ....
....

7. Has any question, asked in an intelligence test, suggested to you a change in method or material at any time? ....

What question? ....

Have you modified method or material accordingly?
....

8. Of your own volition, have you read anything about intelligence tests or their use since February, 1928?
....

If so, what? ....

9. Has your work with intelligence tests led to interest in or taking of courses in measurement? ....

If so, what courses? ....

10. What was your attitude towards tests and measurements before September, 1926? ....

Summary of Teacher Questionnaire

Number of teachers who gave group tests, 1926-1929..70

Q.1. Number of teachers still located in Belmont who gave them 57
Number of Kindergarten and Grade I teachers who gave Binet tests, 1927-1928 21
Number of Kindergarten and Grade I teachers located who gave them 21

Q.2. Number of teachers who report having given Binet tests for their own satisfaction, or to obtain a rating for a pupil who had been absent....... 9
Children thus tested 101
(Also "several" tests by two teachers)

Q.3. Number of teachers who report having of their own accord called special cases to the attention of the Director of Tests, or State Clinic, or Social Service, etc...............................12
Number of cases thus located (This number is verified by Director of Tests)................56
Number of teachers who say they would not have done this if they had not had test experience..12

Q.4. Have made use of information on test record cards (teachers)...27
Uses made by principals: (These principals were testers)

Classification.....................1240 pupils
Grading ....1240 "
Promotion .... 505 "

By teachers:

Study of special cases.......................34
Comparing teacher's judgment with test results (3 teachers "many").........35
Comparing child's work with test results..26
Adapting child's work to child's ability..26
Comparing ability with work produced by child.................................25
To get a basis for judgment (looked up Intelligence Quotients)..............10

Q.5. Teachers who have asked to have group intelligence tests given to their classes since May, 1928... 8

Q.6. Teachers who report having used facts brought to their attention by the test program............15

"What facts" was interpreted as follows:
What teacher did about these facts:
  "Recognized non-readers. Studied the cases further."
  "Drilled on points shown as weak by the tests."
  "Exercised more care in my own judgment of a child."
  "Demanded better work of capable pupils who were not making the most of their ability."
  "Tried to bring out a diffident child who showed unexpected ability in the tests."

Q.7. Teachers not answering at all..........................27
Teachers not finding test material suggestive........26
Binet question suggested the need for more practice in making odd change from a quarter............. 2
Teacher (in science) had occasion to remember that appreciation of likenesses and differences do not usually develop synchronously............... 1
Story of the New York fire (Binet) suggested a method for getting thought-content in reading... 1

Q.8. Teachers reporting readings on the subject of intelligence tests or their use since February, 1928.36
Teachers reporting not having read on this subject...14
Not answering at all................................... 7
"Trying to read all of the new things on the subject." 1
Reporting readings in Educational Journals...........16
The books named were:
  Bronner, Special Abilities and Disabilities..1
  Corning, After Testing--What? 1
  Dickson, Intelligence Tests and School Reorganization.........................1
  Hollingworth, Gifted Children................1
  Kohs, Intelligence Measurement...............1
  Monroe, Measuring the Results of Teaching....1
  Terman, Measurement of Intelligence-- Re-read or studied...................6
  Thom, Child Psychology (pamphlets)...........1
  Whitley, Tests and Scales in Psychology of Childhood.............................1
  Wolfarth, Supervision and Teaching of Reading............................1 15

Teachers reading, but not stating what............... 4

Q.9. Work with tests has led to increased interest........20
Working for A. M. degree in this field............... 1
Taking courses in this field......................... 2
  Normal School Course in Educational Measurement..............................1
  Thom, Child Psychology.......................1
Taking advanced courses in education, not in this field...................................22

Q.10. Teachers reported their attitudes towards tests and measurements before September, 1926, as follows:

Not interested	2	Not exactly in favor	23
Much opposed	6		
Didn't know much about them	13		
Thought them the business of principals and supervisors	2		
Favored the movement	9	Favoring	24
Curiously interested	13		
Interested (but not for Kindergarten)	2		
No answer given	10	Not answering	10
Total	57	Total	57

A check of the 20 teachers who, in answer to Question 9, said that "Work with tests had led to increased interest", shows that all of the 20 are of the "not exactly in favor" group of 23, reported in Question 10.

With substitution of success for failure, with decrease in disciplinary problems, with increased accomplishment, with increase of teacher interest, and a bettering of teacher morale, the small expense of time and labor becomes as nothing in comparison. In Chapter VI it has been shown that expense is more than balanced by income; the expense being $2,409.90, and the saving being $11,301.05.

CHAPTER IX

SUPPORTING DATA

What is the experience of other workers in similar fields?

On April 22, 1929, a questionnaire was sent out to eighty-five persons who are heads of research bureaus or departments in public school systems of the United States, or who have had special experience with work in ability grouping. The questions were those which developed in conference of superintendent, organizer, and director of Tests and Measurements, held near the end of the third year of this study. Sixty-one responses from those whose names, positions, cities, and size of school systems which they served are listed in Table 12, yielded the totals and added suggestions given in this chapter. Those who did not send returns number twenty-four. Five of the twenty-four directors represent systems whose policy is known to be conservative in the matter of answering questionnaires. The other nineteen are either new to their present positions, or are just beginning new organization of research in their school systems.

The fact that one point was checked by a larger number of directors than another does not necessarily mean that that point has more weight with reference to the question,

or more value as a point. It probably does mean that that point has come within the range of experience of more of the directors. Some of the statements given as mere suggestions by one person, may prove to be of more value in certain situations and under certain conditions than the more popular answer. In general the directors expressed their unwillingness to permit a check mark to stand as adequate expression of their answers to these vital questions. This accounts for the fact that sixty-one answers are not given for each question. The numbers recorded are those of actual check marks made. Quotations which embody the leading points brought out in the added discussion of the directors, follow each question.

Letters accompanying the returned questionnaire throw added light upon the problems. Robert Floyd Gray of San Francisco writes: "While your questionnaire raises no end of vital questions relative to the testing program, practically all of the answers would be influenced by the size of the city and the educational policy involved. This is especially true of the questions relative to the responsibilities of the superintendent and of the research director in the testing program. With these factors in mind I have checked the questionnaire."

George R. Johnson of St. Louis says: "My answers to these questions have been determined mainly by three fundamental principles:

(1) Testing is an inseparable function of modern teaching.

(2) The grouping of pupils within a class in an inalienable right of the teacher in charge.

(3) The principal of the school is primarily responsible for the assignment of pupils to the several classes within the school.

"I do not believe that the Director of Research in a school system, or the Superintendent of Schools should interfere with the teacher or principal in the performance of these duties. The Director of Research may render service to the schools in the system by expert counseling, by recommending suitable tests, by collecting and interpreting data, etc. The Director of Research must not, however, usurp the functions of principal or teacher by trying to dictate the internal policies of the school or the classroom."

Harry J. Baker of Detroit writes: "The real problem in regard to classification by 'Ability Groups' is not the matter of actual testing and grouping of pupils, but the adjustment of the course of study to fit the needs of these pupils after the testing has been done. In many cities the details of a testing program requires so much time that nothing is done on follow-up work which is really the excuse for the testing program.

"It seems to me that your chief problem is to sell the idea of curriculum revision and modifications to the administrative officers who have charge of determining the course of study. Up to the present time so much attention has been focused on the mechanics of the testing program that curriculum adjustments have been very much neglected."

The form used for questionnaire follows. Table 12 shows the names of the workers in research who answered the questionnaire, together with their official positions, their cities and the populations of cities and of schools which they serve. Their answers are given in summary from following Table 12.

QUESTIONNAIRE

NOTE:
The answers given are mere suggestions
and are to be marked √ in the right hand
margin if they meet with your approval.
Space is left for any additional answers.

1. What specific data concerning the school system should a director of research have in hand before attempting to plan a testing program upon whose results ability grouping may be based?

- Data from all previous tests ....
- The location and grouping scheme of all of the classes ....
- The districting and housing of pupils ....
- Geographical matters affecting school attendance ... .
- Preparation of teachers for teaching ....
- Amount of training teachers have had in the field of tests and measurements ....
- General attitude of teachers towards a testing program ....

2. (a) To what matters should a superintendent give personal attention during a testing program?

- Assistance of director in planning the program......
- Education of the public with respect to the program ....

- His presence at general teachers' meetings bespeaks coöperation ....
- Any matter which seems to director, teacher or parent to be a source of the least friction....

....

- Expense of the program ....

....

(b) What are the specific duties of a director of research?

- Plan the test program ....
- Plan and conduct teachers' meetings ....
- Be available for conference with teachers on test matters ....
- Direct the testing and classification ....
- Know problem cases as thoroughly as possible ....
- Make oral and written reports to the superintendent upon request ....
- Keep accurate record of procedures ....
- Study findings in the light of the data reported by other workers in similar fields...............
- Make public through writings, findings of value from his research ....
- Conduct further research upon matters suggested by the superintendent ....

....

....

(c) What matters demand conference of superintendent and director?

- Points of question or friction in carrying out plans ....
- Parental objections to the classification of children ....
- Unresponsiveness of a teacher ....
- Matters of policy arising from classification:
  - Changing of teacher's grade or room ....
  - Changing of pupils from one building to another ....
  - Establishing of special classes or groups......
  - Undertaking an experiment, such as a Pre-School ....
  - ....

(d) Assuming that intelligence tests to be administered by teachers, what is the share of the school principal in a testing program for classification?

- Distribution and collection of materials ....
- See that summaries, reports, and records are prepared correctly, and at the specified time ....
- Attend to the actual changing of all pupils from one group to another ....
- Keep office records of the classification in form at all times ....

- Meet questions of parents ....
- Know and note cases needing special study and diagnosis ....
  ....

3. (a) During the first year of a testing program how much time is needed for general and group teachers' meetings?

- One general meeting before, and one or two after each test period ....
  ....
- Group meetings enough to develop thorough understanding of all of the details which the teacher must meet in carrying on tests ....
  ....

(b) What matters should be discussed at such meetings?

- At general meetings, purposes, and objectives ......
  ....
- At group meetings:
  - Meaning of each part of the test to be given ..
  - Method of giving the test ....
  - Method of scoring and summarizing ....
  - Directions as to time and place of returns, etc...
  - Returns, with possible interpretations ....
  ....

4. How much actual administration of Stanford-Binet Tests should a teacher have done successfully before she may be entrusted to do this testing accurately without supervision? (Assume that the teacher is interested enough in the work to have undertaken it voluntarily.)

- 10 tests ....
- 30 tests ....
- Some teachers ought never to be entrusted to give intelligence tests ....

....

5. Is it best that the director of research administer special tests for problem cases?

- Yes, if the school system has not or cannot get coöperation of a clinic ....
- No. Reasons:

....

....

6. (a) Are there measures of teachers' errors in giving and scoring group tests of intelligence?

....

....

(b) Does the clerical work of teachers in checking and recording test results give returns in better understandings?

- The mechanical part of this work is apt to be uppermost in the teacher's mind ....

- Yes. Each check is bound to give insight into the workings of the particular test checked........

....

....

7. How much clerical service per 500-pupil unit is needed for a year's testing program?

- Half time of one clerk.................................
- Full time of one clerk.................................
- Usual clerical service is sufficient to handle this part-time work...............................

....

....

8. (a) If teachers' meetings have been managed wisely, how much adjustment should be needed between teachers upon questions which arise from changes in classification?

- A great deal...

....

- None..

....

(b) What are these questions?

- Which teacher shall take the low division? ....
- Is it fair for a teacher to give up some of her best pupils, and take in exchange the same number from a low division, for the sake of classification? ....

....

....

9. (a) Should superintendent, director of research, or principal meet the questions of parents?

- Superintendent, because his advice will be accepted more readily by certain parents................

....

....

- Director of research, because he has the test records, and can advise authoritatively........

....

....

- Principal, because he knows parents and pupils best.

....

....

(b) How much time is this person justified in giving to conferences with parents upon matters brought out by the testing and classification?

- As little as possible...............................
- A special office hour is justifiable for this purpose..
- Every parent has a right to as much time as it takes to bring about a satisfactory understanding....

....

....

....

10. Who should handle remedial work with parents who need the assistance of mental hygiene?

- A clinic worker, if available..........................
- School nurse or visitor, or teacher.................
- The person who has promise of being able to understand and influence the parent.................
- Only a worker skilled in mental hygiene............
- This is not the province of the school..............

....

....

11. Should a superintendent's program for education of the public parallel or follow a testing program?

- Parallel, that the public may not have apprehension because of misunderstandings...................

....

....

- Follow, because public addresses and reports may then be framed to meet the questions which have actually arisen in the given situation.........

....

....

....

Table 12

Leaders in Public School Research

	Names	Positions	Cities	Estimated Pop. of City	Approx. Pub.Sch.Pop.
1.	Richard D. Allen	Asst. Supt. in Charge of Research and Guidance	Providence, R. I.	286,000	41,000
2.	Harry J. Baker	Clinical Psychologist for Public Schools	Detroit, Mich.	1,500,000	240,000
3.	F. H. Beede	Superintendent of Schools	New Haven, Conn.	184,900	30,000
4.	H. H. Bixler	Director of Tests and Guidance	Atlanta, Ga.	250,000	62,000
5.	Philip Boyer	Director Division of Educational Research	Philadelphia, Penna.	2,000,000	250,000
6.	Ernest P. Branson	Director of Research, Public Schools	Long Beach, Calif.	155,000	26,000
7.	Wesley L. Brown	Director of Research	Winnetka, Ill.	30,000	1,700 High School
8.	Harold B. Chapman	Director of Educational Research	Baltimore, Md.	825,000	162,000
9.	C. L. Cushman	Director of Research	Oklahoma City, Okla.	180,000	30,000
10.	Prudence Cutright	Asst. Director of Instructional Research	Minneapolis, Minn.	447,700	90,000
11.	William L. Connor	Director of Bureau of Educational Research	Cleveland, Ohio	1,200,000	150,000
12.	Murray Dalman	Director of Educational Research	Indianapolis, Ind.	374,300	75,000
13.	Charles D. Dawson	Director of Research	Grand Rapids, Mich	180,000	28,000
14.	Virgil E. Dickson	Director Bureau of Research	Berkeley, Calif.	69,400	13,000
15.	Russell Doney	Director of Research	Kalamazoo, Mich.	60,000	10,000
16.	Gertrude B. Duffy	Director Department of Measurement	Cambridge, Mass.	120,000	17,000
17.	R. R. Erwine	Assistant Superintendent	Akron, Ohio	240,000	65,000
18.	L. L. Everly	Director of Research	St. Paul, Minn.	250,000	38,000
19.	Violet Foster	Director of Special Classes	Jackson, Mich.	61,700	10,000
20.	Guy Fox	Assistant Director of Research	Denver, Colo.	256,491	47,434
21.	Robert F. Gray	Director of Research, Public Schools	San Francisco, Calif.	756,188	85,122
22.	R. L. Herbst	Director of Tests and Measurements	Lancaster, Penna.	58,000	9,000

(Table 12 continued)

	Names	Positions	Cities	Estimated Pop. of City	Approx. Pub.Sch.Pop.
23.	J. G. Hickox	Director of Research Department, Public Schools	Warren, Ohio	45,000	8,000
24.	Harriet Hinnon	Director of Research in Public Schools	Toledo, Ohio	325,000	61,233
25.	W. Hardin Hughes	Director Bureau of Research	Pasadena, Calif.	69,500	13,000
26.	Clara Jacobs	Director of Research, Public Schools	Pueblo, Colo.	65,000	12,000
27.	George R. Johnson	Director of Division of Tests and Measurements	St. Louis, Mo.	945,000	100,000
28.	Arthur W. Kallom	First Assistant Director, Department of Instructional and Educational Research	Boston, Mass.	900,000	165,975
29.	Grayson H. Kefauver	Director Bureau of Research	Fresno, Calif.	62,200	11,000
30.	A. L. Kerbow	Director of Research	Houston, Tex.	300,000	60,000
31.	Ruth Kimball	Director of Testing and Research	New Britain, Conn.	85,000	13,000
32.	Frank H. Koos	Director of Research	Winston-Salem, N. C.	70,000	13,000
33.	Helen M. Kopf	Director of Research	Piqua, Ohio	17,500	3,000
34.	Edward A. Lincoln	Professor of Education, Harvard University	Cambridge, Mass.	120,000	17,000
35.	Ida C. Lucas	Psychologist, Department of Research and Guidance	Reading, Mass.	10,000	2,000
36.	L. D. Lundberg	Director Department of Research and Statistics	Flint, Mich.	150,000	30,753
37.	Clara Mallory	Director of Research	Beaumont, Tex.	70,000	8,506
38.	Arthur L. Maxon	Director of Research	Schenectady, N. Y.	95,000	20,000
39.	Joseph M. McCallie	Asst. Supt. New Jersey State Home for Boys	Trenton, N. J.	140,000	22,000
40.	Eugene E. McCarthy, Jr.	Director of Elementary Instruction	Dade Co., Miami, Fla.	140,000	25,000
41.	Elizabeth McFadden	Director of Research	Colorado Springs, Colo.	37,000	12,558
42.	George W. Melcher	Director of Educational Research	Kansas City, Mo.	383,100	65,000
43.	George O. Moore	Assistant Superintendent	Erie, Penna.	130,000	25,500
44.	H. B. Nash	Director of Research	West Allis, Wis.	32,000	5,773
45.	C. H. Nettels	Assistant Director of Research	Los Angeles, Calif.	1,300,000	200,000
46.	Eugene A. Nifenecker	Director of Bureau of Educational Research	New York City	6,064,484	1,028,671
47.	W. J. Osburn	State of Ohio, Dept. of Education	Columbus, Ohio	291,400	50,000

(Table 12 concluded)

	Names	Positions	Cities	Estimated Pop. of City	Approx. Pub.Sch.Pop.
48.	Charles Ottermann	Assistant Superintendent	Cincinnati, Ohio	420,000	98,899
49.	Mary M. Pfefferkarn	Director of Research, Public Schools	Leavenworth, Kans.	22,000	2,500
50.	Hattie O. Prewett	Director of Research, Public Schools	Fort Worth, Tex.	200,000	33,234
51.	Henrietta Race	Psychologist, State Dept. of Education	Madison, Wis.	48,800	8,500
52.	J. L. Stenquist	Director of Bureau of Research	Baltimore, Md.	825,000	162,000
53.	David R. Sumstine	Director of Research in Public Schools	Pittsburgh, Penna.	666,000	110,000
54.	Katherine A. Tilton	Director of Psychological Research	San Diego, Calif.	150,000	26,000
55.	Betty W. Tracy	Director of Special Education	Wheeling, W. Va.	56,200	9,000
56.	Eston V. Tubbs	Director of Curriculum	Chicago, Ill.	3,000,000	500,000
57.	Charles W. Walker	Director of Tests and Measurements	Concord, N. H.	25,000	3,300
58.	Ralph W. Walter	Director of Tests and Measurements	New Rochelle, N. Y.	50,000	8,500
59.	Roscoe L. West	Director of Teacher Training, State of New Jersey	Trenton, N. J.	140,000	22,000
60.	Ernest C. Witham	Director of Research	Wilmington, Del.	128,000	16,516
61.	J. T. Worlton	Director of Research	Salt Lake City, Utah	135,000	32,000

Summary of the Sixty-one Responses
to the Nation-wide
Questionnaire

Question 1. What specific data concerning the school system should a director of research have in hand before attempting to plan a testing program upon whose results ability grouping may be based?

- "Location and grouping scheme of all the classes" was checked by 46
- "Data from all previous tests" by 42
  - Data from recent tests 3
  - Data from related tests 1
- "Amount of training teachers have had in the field of tests and measurements" 42

Tilton adds, "In actual practice, not in compulsory courses in Training Schools."

Lincoln: "It is especially important that teachers be trained in teaching the different ability groups."

- "General attitude of teachers towards a testing program" 42

Kimball: "Much depends upon the mental calibre of the teachers."

Duffy: "Consider the amount of clerical work already incumbent upon teachers."

Allen: "This point is very important (teacher attidude). Go slowly. Digest each test before trying another."

Johnson: "Data showing need of the program will help this."

Everly: "The right attitude can be developed by first making tests optional. 'Sell' the idea through the advantages obtained."

- "Districting and housing of pupils" 32
- "Geographical matters affecting school attendance" 27
- "Preparation of teachers for teaching" 25

Gray and Kallom state that this may be a factor in the success of the program, but it is not a responsibility of the research director.

Question 2. (a) To what matters should a superintendent give personal attention during a testing program?

- "Expense of the program"("to justify it")........ 42

Only in the sense that any educational feature is the superintendent's concern..... 1

Tilton: "Curtailing expense is bad, when teachers coöperate and ask for plenty of material."

Everly: "The amount of money needed should be secured by placing it in the budget."

Lincoln: "As the work proceeds the superintendent should need to give less and less attention to expense."

-"Assist the director in planning the program 36

The superintendent should approve only........... 2

Herbst: "It is the director's job to plan according to suggestions by the superintendent."

- "His presence at general teachers' meetings bespeaks coöperation"........................ 36

However, it is not necessary................. 4

This depends much upon size of system....... 1

- "Education of the public with respect to the program" 35

If this is necessary......................... 1

- "Any matter which seems to director, teacher, or parent to be a source of friction".......... 23

If it cannot be settled without him......... 1

(b) "What are the specific duties of a director of research?"

Tilton: "These vary with the superintendent's policy, with his scope of vision in detailing responsibilities to members of his supervising staff."

Moore: "The director's department should be a fact-finding agency only."

- "Keep accurate records of procedures" 54
- "Be available for conference on test matters".... 53
  - "Especially with principals" 1
- "Plan the test program" 52
- "Make oral and written reports to the superintendent upon request" 52
  - "Without request" 1
- "Conduct further research upon matters suggested by the superintendent" 52
  - Include assistant superintendents in a large system 1

Lincoln: "Matters for research should be contributed also by interested teachers and principals."

Moore: "Policies and dealings with the Board come through the superintendent. He determines policies and presents them to the Board in the light of facts furnished."

- "Study findings in the light of the data reported by other workers in similar fields" 51
- "Direct the testing and classification" 50
  - Yes, through the principals 1
- "Know problem cases as thoroughly as possible"... 47
  - In a small system, yes 1

Dawson, Gray, and Stenquist: "Not personally if the system is too large."

- Make public through writings, findings of value from his research	43
"If the superintendent approves"	1
"Reports should be through the Board of the city"	1
"Reports should be at the disposal of teachers and principals"	1
"Not always"	1
- "Plan and conduct teachers' meetings"	38
"When necessary"	1
"Occasionally"	1
"Should assist at the meetings"	1
"This is the principal's responsibility"	1

Additional suggestions offered as falling within the director's province:

Witham: "The director should be a tactful free lance. He should work for the good of the system, not to show off."

Osburn: "Ideally the director does his job so well that the superintendent does not need to suggest."

Stenquist: "He should handle conferences over results."

Herbst: "He should plan remedial measures with teachers and principals."

Nash: "He should direct construction of objective tests, correlating abjective testing

and teachers' marks."

Walter: The director's concern is also "with curriculum distinctions for differentiated ability groups."

Foster: "In small systems his duties may include supervision of special classes, also education of teachers and community along mental hygiene lines."

(c) "What matters demand conference of superintendent and director?

- Points of question or friction in carrying out plans	50
"There should be few"	1
- "Matters of policy arising from classification:	
"Establishing of special classes or groups".	38
"Undertaking an experiment such as a Pre-School".	37
"If done"	1
- "Changing of pupils from one building to another"	35
This is an administrative problem	1
No	1
- "Changing of teacher's grade or room"	24
This is the principal's business	3
An administrative problem	2
Often handled by principal and director	1
Not usually a matter for conference	1

- "Unresponsiveness of a teacher"	31
This is a principal's duty	5
Question this	2
As a last resort	1
Depends upon extent of the trouble	1
- "Parental objection to the classification of children"	26
Seldom met	3
Never met	2
Met by director	1

Added suggestions:

Fox and Nifenecker: "Size of system plays an important part, and the duties of assistant superintendents are concerned."

McCallie: "Don't begin wholesale Pre-School or any other experiment. Try out in a school or two where conditions are favorable to success."

Moore: "Pre-School is hardly fact-finding. This is done through use of Parent-Teachers' Association."

(d) Assuming that intelligence tests are to be administered by teachers, what is the share of the school principal in a testing program for classification?

- "See that summaries, reports, and records are prepared correctly, and at the specified time" 53
- "Keep office records of the classification in form at all times" 53
- "Meet questions of parents" 52

  Or refer to director 1
- "Know and note cases needing special study and diagnosis" 51

  Lincoln: "In some cases it might be better to select a teacher in each building for this, especially if there is a teacher better trained than the principal in educational measurements."
- "Attend to actual changing of pupils from one group to another." 47
- Distribution and collection of materials 42

  No 1

  If there is no other way 1

  Kallom: "See that reports of the director reach the teachers who are concerned in the report."

  Added notations:

  Baker: "We have trained examiners in a large system."

  Worlton: "Teacher administration of psychological tests is a poor practice."

Allen: "We do not believe in teacher administration of psychological tests."

Nifenecker: "Our principals supervise and train teachers to test."

Gray: "San Francisco plans a 'testing adviser' for each school for special study cases."

Stenquist reports that Baltimore is working out a system of test administration whereby certain representative teachers, trained in testing, are responsible for administration of tests and for returns, and receive added salary for successful service. (167: 8) also (168: 1-11)

Question 3. During a first year of a testing program how much time is needed for general and group teachers' meetings?

- "Group meetings enough to develop thorough understanding of all of the details which the teachers must meet in carrying on tests" ... 47

Johnson: "Usually clearly expressed written directions will suffice."

Lundberg: "Supplement with mimeographed directions and 'don'ts'."

Tracy: "Special training in giving and scoring tests comes in here, also practice in timing, etc."

Tilton: "Avoid meetings. The practice breaks down interest and coöperation."

Moore: "Such meetings are not effective."

- "One general meeting before and one or two after a testing period" 32

McCallie: "This depends upon training of teachers."

Allen: "One before and one after."

Nash: "Principals take care of this."

Stenquist: "We spent the major part of the first three years 'selling' the program."

(b) What matters should be discussed at such meetings?

- "At general meetings":

  Purposes and objectives 21

  These can best be shown by principals 19

Lincoln: "After testing, general results, plans for remedial measures, etc."

- "At group meetings":

  Method of giving the test 48

  Method of scoring and summarizing........... 48

  Directions as to time and place of returns.. 44

    This should be by bulletin 3

  Returns with interpretations 38

  Lucas: "Very little of this."

  Kallom: "After the test has been given and

re-checked, the returns with interpretations should be discussed in small groups, and with individual teachers."

- Meaning of each part of the test to be given 36

Question necessity of this 2

Added suggestions:

Witham and Gray: "Give mimeographed directions in addition to oral ones, and give much of the returns in mimeographed form."

Boyer: "We give written standard directions."

Moore: "Elementary teachers resent too many meetings. They will come through loyally, but we have found finer coöperation through carefully prepared mimeographed sheets."

Connor: "Be sure to distinguish between clerical work and interpretation. The teacher should be taught to do the clerical work mechanically and rapidly. She should study interpretation afterwards."

Kimball: "Be sure to stress with teachers the need for accuracy and adherence to directions without changing them."

Johnson: "This idea of returns is all right if practicable, the purpose being to secure exchange of ideas."

Question 4. How much actual administration of Stanford-Binet Tests should a teacher have done successfully before she may be entrusted to do this testing accurately without supervision? (Assume that the teacher is interested enough in the work to have undertaken it voluntarily.)

There was general expression of unwillingness to check a definite number of tests. However,

"Some teachers ought never to be entrusted to
give intelligence tests" was checked by 43
"30 tests" by 33
"10 tests" by 9

All stress the need of the right teacher for a tester, and of the necessity of courses or other "proper background" in psychology before taking up the study of intelligence testing.

Connor: "The giving of a Binet Test is a comparatively simple affair. The interpretation of the test is a very different matter and requires a person of peculiar abilities and rather long experience."

Baker and Maxon: "We need full time, experienced, trained examiners."

Osburn: "Practically some teachers ought never to be entrusted to give intelligence tests. Theoretically, any real teacher should be able to do it."

Johnson: "Some teachers ought not to be entrusted to give tests --- or to continue the work of instruction; but these cases should have individual administrative consideration."

Foster: "It is possible to train kindergarten teachers to give tests to pupils for Grade I B with from 10 to 30 tests. This would not be enough for long range testing."

Question 5. Is it best that the director of research administer tests for problem cases?

All who answered stressed the time factor, the size of system and of staff, and ability or non-ability or lack of training of the director to do this specialized work.

Yes, if clinic service is not available	45
No	6

Baker: "The director should probably be the head of the Psychological Clinic."

Everly: "Special cases should be handled by specialists."

Kallom: "Clinic coöperation is needed particularly in cases involving mental hygiene."

Question 6. (a) Are there measures of teachers' errors in giving and scoring group tests of intelligence?

"I know of none"	6
"Have no data"	6
"Have data, not yet compiled (Everly)	1

Connor: "P. E. of Probable Learning Rate from group mental tests in Cleveland is now known to be 7."

Allen and Nifenecker: "All tests should be re-checked."

Dickson: "Re-testing is about the only method available."

Foster: "Teachers' work, errors, and attitude make it best to entrust the scoring to experienced clerks."

Kallom: "There are measures which pretend to determine the amount of errors in any given group which has been tested. They need not be reported to teachers at all, as most teachers have not had the training which would make it possible for them to understand the report."

Lincoln: "There are studies which show that even skilled workers make errors. And I sometimes make a slip in giving a test myself."

Gray: "Error will always be a factor, but not enough to invalidate all results."

(b) Does the clerical work of teachers in scoring and recording test results give returns in better understandings?

- Mechanical part is apt to be uppermost in the teacher's mind 31

Not to a real teacher 3

"Depends upon the type of teacher" Hinnan-Nifenecker 2

Jacobs: "It will be uppermost until the teacher realizes the help the test papers are to her in understanding the child's reaction."

Connor: "It should be mechanical to insure rapid and accurate scoring."

Baker: "It should all be done by clerical help."

- Yes. Each check is bound to give insight into the workings of the particular test checked 20

Brown: "Yes, for a few tests."

Kimball: "Amount of insight depends upon teacher's interest in and knowledge of tests."

Everly: "It is true only for better trained teachers."

Herbst: "Yes,- insight into pupil difficulties also."

Four, reporting for cities in which intelligence tests are given by director's department, and educational tests by teachers and principals, write, "Yes, on educational tests."

Foster: "Too many demands are made upon teachers to have it operate that way."

Tracy: "It will give teacher insight, if done with that specific purpose."

Fox: "Probably not for most teachers. Study of checked papers will give greater benefit."

Kallom: "The testimony of teachers varies a great deal, from the teacher who says she gets nothing and it is only a burden, to the teacher who gets a great deal out of the correcting. It is apparently determined by the interest which the teacher has in the use of the final results."

Question 7. How much clerical service is needed per 500-pupil unit for a year's testing program?

-"Usual clerical service is sufficient to handle this part-time work"	23
Depends upon what "usual clerical help" is	1
-"Half time of one clerk"	12
-"Depends upon extent of the program"	5
-"Full time of one clerk"	4
-"Do not know"	1

Nifenecker: "Machines tabulate results in New York, so I cannot answer."

Maxon: "This depends upon the amount of clerical work made by the director."

Moore and Lincoln: "It depends upon the tests to be given, and upon what is to be done with results."

Bixler: "It depends upon the amount of work which teachers do."

Allen: "500 tests means 50 hours of work."

Baker: "7 clerks have scored 95,000 tests in 7 months."

Connor: "1 clerk can score 500 tests in 2 months,
or
4 clerks can do the same work in 2 weeks."

Foster: "1 clerk can score for 7,000 children in 3 or 4 months."

Tracy: "1 clerk handles work for 2,000 when tests have been scored by the teacher."

Walter: "1 clerical assistant for 8,000 children."

Question 8. (a) If teachers' meetings have been managed wisely, how much adjustment should be needed between teachers upon questions which arise from changes in classification?

- None, if principal is supervisor as well as administrator	18
- Very little	12
- Some	11
- A great deal	4

Fox: "This depends upon the teachers."

Allen: "Some teachers never will understand."

Gray: "This depends upon the matter of principal's supervision."

Everly: "Depends upon (1) the amount of teacher's training and experience, and (2) the amount of detailed and lucid interpretation which is given by the principals and the Bureau of Research."

(b) "What are these questions" (having to do with adjustments between teachers)?

- Which teacher shall take the low division? 19

  The best qualified teacher should take the low division............................... 4

Tubbs: "The best teacher should take the low group."

Tilton: "The teacher with the social view point should take the low division."

Kimball: "A trained teacher is most desirable for the low division."

Nettles: "Teacher personality must be studied. Few teachers should have all low groups."

Foster: "Some principals alternate teachers."

Kallom: "I am rather certain that the difficulty of getting teachers to handle the slow section is that in spite of the fact that the principal knows it is a slow section, because of pressure from supervisors and superintendents *he makes the teacher feel that she is to get as good results* in her class as he is getting in the other classes in the school. Where you have a *real* special class you have teachers who are perfectly willing to deal with these children who are low mentally but in no case is there pressure brought to bear by superintendents, directors, or principals to bring about certain preconceived results comparable to the work in

any particular grade."

Kefauver: "Such questions should not arise among teachers."

Kerbow: "We have no trouble here."

- Is it fair for a teacher to give up some of her best pupils, and take in exchange the same number from a low division, for the sake of classification?

Yes	38
This is an administrative matter	16
No	1

Johnson: "It is fair if the policy is good."

Stenquist and Dickson do not see this as a matter of "fairness". "The teacher should consent."

Kallom: "If it is best for the pupils to go into a brighter section, there is no question at all of being fair to the teacher. I am very certain in my own mind that keeping bright children in sections with slow children in order to furnish a stimulus to the dull children does not work. In running a race we do not put the pacer so far ahead that it is impossible for the other parties in the race to catch up. I have also seen large numbers of cases where children placed in slow sections have for the first time been happy in their school work."

Gray: "A plan of homogeneous grouping should eliminate this."

Everly: "It is fair, but teachers and principals do not always see this matter professionally."

McCallie:- "Pupils' good, not teachers' pleasure is the consideration."

Connor: "Later, teachers will learn to avoid bright sections unless especially adapted for the work."

Lincoln: "Follow teachers' special interests in this. If teachers aren't already in this work who will do it willingly and eagerly, others should be found."

Further questions suggested:

Fox: -"Techniques for handling different divisions.
- Modifications in course of study for high and low.
- Securing of supplementary or different materials for different groups."

Prewett: -"The relative number of pupils in A, B, and C groups enters here. A teacher in a low division should have fewer pupils."

Question 9. (a) Should superintendent, director of research, or principal meet the questions of parents?

The question was answered by all.

Though the phraseology of answers varied, the thought of 59 was identical; namely,

First appeal of the parent should be to the principal. If necessary, and sometimes it is most advisable, second appeal should be to the director for his facts. The superintendent should be the "court of last resort". "Each must do his part, and all must coöperate in the best interests of the child."

Two would make the order of reference, principal, superintendent, director. This variation may be due to other conditions in these two particular school systems which have to do with duties and responsibilities of superintendent and director.

Maxon: "The director is a staff officer, not a line officer concerned in execution of details."

(b) How much time is this person justified in giving to conferences with parents upon matters brought out by the testing and classification?

- "Every parent has a right to as much time as it takes to bring about a satisfactory understanding". 43
- "A special office hour is justifiable for this purpose".. 9
- "As little as possible" 8
- "A special office hour is not justifiable" 1

Baker: "This is no problem at all. Measurement is an internal school device."

Gray: "A testing program properly handled will result in little misunderstanding."

Nifenecker: "A reasonable amount should be given each parent."

Foster: "Time is well spent explaining these things to parents when they inquire. It is to them that we must ultimately justify ourselves."

Lincoln: "This will pay in the long run, because of the part it will play in the education of the general public."

Question 10. Who should handle remedial work with parents who need the assistance of mental hygiene?

The question was generally interpreted as meaning "Who should approach the parent when the child needs mental hygiene help?" It was the author's intention to ask "Who can best handle the parent problem, when the parent himself is known by the school authorities to be an unsatisfactory factor in the child's reactions?"

Herbst: "If the parent is meant, then it is not a school problem."

Melcher: "The school should help if it affects the welfare of the child."

As best approach to mental hygiene help for the child,

- "A clinic worker, if available" was checked by ..	29
- "Only a worker skilled in mental hygiene"	18
- "The person who has promise of being able to influence the parent"	16
- "School nurse"	15

Walter: "A visiting teacher can often do much."

Dawson: "The person who is most capable of understanding the child."

Kimball: "Coöperation of all who can."

Kallom: "Only a clinical worker in mental hygiene, a psychologist, or psychiatrist should handle the remedial work with such parents. School nurses, school visitors, teachers, and principals have neither training nor information to handle such cases. This skilled worker should have the coöperation of everybody who is dealing directly with the child."

Question 11. Should a superintendent's program for education of the public parallel or follow a testing program?

- Parallel	19
- No program unless the public demands it	16
- Follow	11
- Parallel	11
- Precede, parallel, and follow	4

Everly: "Even in a first testing program, the superintendent's educational influence needs to be directed first to the Board, the principals, and the teachers."

Gray: "The testing program should not be presented as an innovation, but as an aid in making the educational program more effective."

Baker: "Do nothing unless you seek trouble."

Sumstine: "The less said, the better."

Stenquist: "No two will be apt to follow identical procedure."

Lincoln: "Probably a judicious combination of both is desirable."

CHAPTER X

SUMMARY

This study was instigated by a situation which developed in a school system of approximately 3,700 pupils through incomplete and wrong use of tests, and through unsatisfactory handling of the resulting material. Up to September, 1926, the test work in Belmont, Massachusetts, had been sporadic, and carried on by teachers according to uncoördinated or no plan. No definite program had been laid out, nor even main objectives known. The results were three: first, a superintendent who was puzzled as to what the next move should be; second, wearied principals and teachers who criticised and discussed and took sides as to values of a measurement movement; third, time, labor, and money were wasted upon tests and test materials which yielded no returns to the children and which really did develop wrong teacher attitudes. A better classification was sadly needed in the schools where the gifted, unrecognized as such, were lagging, and the slow were unfairly pushed beyond their ability to function. This situation was believed to be not at all unusual in school systems in which superintendent and teachers, with the best of intentions and with most commendable desire to modernize their schools, have attempted to work with tests without skilled guidance and without the necessary training and knowledge.

The main problem of the study was that of discovering whether it is possible in such a situation to accomplish a program of homogeneous grouping with coöperation of and known benefits to teachers, with coöperation of and understanding by parents, with benefits to and without detrimental effects upon children.

In the attempt at solution of the main problem, two secondary problems developed: (a) What are the problems which arise in a program looking towards homogeneous grouping of pupils in a school system? (b) What is the cost of such a program in time, labor, and money? What are the compensations for the cost?

The evidence has been obtained through three years of actual work upon a carefully planned testing program, beginning with the use of intelligence tests, administered by teachers, and under the direction of a research worker who is a trained organizer of measurement. Problems were recognized as they arose in the actual field, and solutions and values were sought and found. The problems were studied in the light of the best information and experience available. They were acted upon as the data dictated. This is in accordance with McCall's method for determining the truth: "to seek authority, to speculate as to possibilities, to experiment or measure, to accept facts shown by the data, and to act in the light of these facts". (65: 1)

Securing of coöperation and understanding of teachers in the development of a test program, involves right objec-

tives, expert planning, selection of right tests, and wise guidance of teachers in their coöperative work of testing and interpreting results.

The task of leadership is not for a novice, and it should not be tampered with by novices. Objectives must be right. Plans must be expertly made. Tests must be rightly chosen. Guidance must be wise.

Teachers can be trained quickly to administer group intelligence tests, if the teachers are fairly intelligent, and are guided towards an open-minded attitude towards the testing and its resulting problems. Method of training is that of helping the teacher to understand each move in the testing, and of giving opportunity for her to become skilful in its execution.

The general plan, as it grew, included:

- Definition of duties, and field and extent of authority of the organizer, and of school officials concerned with the program.
- Study of any possible values from past testing.
- Detailed plan for a survey to be made by the use of teacher-administered tests.
- Laying out of the objectives of the program in such a way as to shift teacher interest from the test itself to the child and his reactions.
- Planning for ample opportunity for teacher-participation in open discussion of and questioning about the program.
- Training of the teachers to give, score and summarize the particular tests used for their classes. This training of the teacher upon the actual task in hand was a most valuable part of the undertaking.

- Arrangement for obtaining the necessary clerical service, as need for the same became apparent.
- Plans for necessary record forms and files for preserving the data, and making ~~it~~ them accessible.

Tests were chosen because they met the known criteria of good tests, especially in the matter of standardization, and correlation with the Stanford Binet tests.

Teachers were brought into the plan as testers and recorders. In doing the testing and scoring, teachers recognize certain individual differences which make for better understanding of the child, provided the teachers have been properly helped and guided in the work.

The training of teachers to do the testing included:

- Taking of the test as given by an expert.
- Discussion, questioning, and review of procedure.
- Practice in giving the test, assisted by criticisms of the organizer.
- Practice in making class summaries.
- Giving of the test to their classes.

The training upon a test may well be given immediately preceding the time for its administration in order that details may be properly remembered, and that the work may be done in the spirit of the meeting at which training is given.

The first testing will show: (1) Need of developing a form for return of data to the teacher. The forms used in Belmont are shown on pages 43-45. The purpose of this particular form was to center the teacher's attention upon

each child's position within his group, and upon the groups to be found in her room; (2) A need for office record form which will preserve the data in a simple, easily understood manner. The material so preserved is shown in Table 4, page 52; (3) Need for a scheme of classification. The data from this study indicated as a working scheme that age might be used for grade placement, and IQ for grouping within a grade. The right placing of individuals was determined in the light of available data as to health, conduct and scholarship. Classifications were not made immediately. Time was allowed for teacher understanding to grow. Except for eight special cases no classifications were made in Belmont until the opening of the second year of the program. Right teacher-attitude and interest in individual pupils cares for the necessary elasticity of classification, provided the administration has the same right attitudes and interest.

All of the teachers and pupils of a school system can be included in the program of testing and classification to the advantage of all concerned. This can be accomplished as rapidly as the teachers can come to accept the newer aims of education and can realize values to be obtained from objective measures, enough to wish to share in the use of tests and in their results. This teacher growth extends over a period of years. In Belmont all of the elementary school teachers and pupils of grades below Junior High School were involved in the program during the first

year, and the group entering Junior High School was classified. The end of the second year found the elementary grade teachers and pupils involved in the classification and benefiting from its results, and two grades in Junior High School classified. The third year showed increase in interest and coöperation of all teachers below Senior High School, with Senior High School pupils tested and with evidence of increase of Senior High School teacher interest.

After first classification, needs should be recognized and studied, and response of administrative force to most urgent needs should be prompt. The need for constant readjustment of classifications should be recognized and met. Urgent needs will press until friction threatens if they are not met. It is far more feasible to meet the need than to remove friction and its causes.

School entrance age is the age for intensive first study. There is no danger of testing intelligence too early, nor in making adjustments for ability in the light of such testing. Proper recognition of ability at school entrance can prevent retardation, and can spare the high school the burden of problems from retardation which are school-life-old.

Skilled assistance is needed for the direction of tests and measurements. The expense of skilled leadership and direction is easily justified. Actual years of repeating done by pupils registered in the Belmont schools on May 1, 1929, had cost the town $39,921.71. The cost

of actual repeating in the school system, cost figured on per capita basis, for the two years prior to 1926, was $13,465.08. In the two years following the beginning of the test program, 1926-1927, and 1927-1928, the cost of retardation was $6,084.30, even though the number of pupils had increased. This was an actual saving to Belmont of $7,380.78, a tidy salary for a director, as well as ample funds for materials for the program. Such classification is a money making proposition. The savings from proper placing of accelerates during the study was $3,920.27. The investment of $.50 per pupil yielded a money-gain of $2.40 per pupil to Belmont.

Problems of a new worker in the classification program for the first year are those of making adjustments in such a way as to permit the good of the former regime to continue uninterruptedly, while loose ends are caught up and the way made smooth for later plans and work.

Blunders can be avoided if the test program is carefully planned by a skilled leader. A program cannot be successful if hurried. For growth, periods of need and adjustment must alternate, and time is a requisite for learning. Early, well planned conferences are a means of forestalling blunders as well as saving time.

The cost in time of a test program for Belmont for three years was almost 82 six-hour days of work by the organizer, who was an unpaid research worker; one year of full time work of a trained director of tests and measurements (salary

$2000); almost 22 hours of the superintendent's time in conference; and $409.90 for test materials. This is decidedly overbalanced by the total saving of $11,301.05. Belmont's clear profit in money in three years was $8,891.15.

The smaller part of the profit is thus measurable in dollars and cents. The substitution of success for failure, the decrease in disciplinary problems, the increase in accomplishment, the increase in teacher interest, the bettering of teacher morale, are vast profits as yet inaccurately measurable. A program for success is not easily nor hastily accomplished. It takes time, thought, teaching, study, wise planning, competent direction and guidance over a period of years. The investment is in conservation of character and contribution to life of the school child -- values which are easily sensed by a teaching staff who have become reasonably keen upon the value of the individual child.

BIBLIOGRAPHY

----

1. Allen, Dichard D. "The Teacher and the New Testing Program." Providence Public School Bulletin, I: 1, March, 1926.

2. ----------------- "A Study in Age-Grade Progress in our Schools." Providence Public School Bulletin, I: 2, June, 1926.

3. ----------------- "The Teacher and Coöperative Research." Providence Public School Bulletin, II: 4, November, 1927.

4. ----------------- "'Learning' the Children." Providence Public School Bulletin, III: 3, March, 1928.

5. ----------------- "Why Teachers Should be Interested in what Becomes of the Product of the Schools." Providence Public School Bulletin, IV: 2, May, 1929.

6. Bailor, Edwin M. Content and form in Tests of Intelligence. Bureau of Publications, Teachers College, Columbia University, 1924.

7. Baker, Harry J. Characteristic Differences in Bright and Dull Pupils. Public School Publishing Co., 1927.

8. Bagley, William C. "Educational Determinism, or Democracy and the IQ." School and Society, XV, April 8, 1922.

9. ----------------- "Educational Determinism Again." School and Society, XVI, August 5, 1922.

10. ----------------- Determinism in Education. Warwick and York, 1925.

11. Bere, May A Comparative Study of the Mental Capacity of Children of Foreign Parentage. Bureau of Publications, Teachers College, Columbia University, 1924.

12. Binet, Alfred "Attention and Adaptation." L' Annee psychologique, VI, 1899. pp. 248-404.

13. Binet, Alfred and Simon, Th. "Methodes nouvelles pour le diagnostique du niveau intellectuel des anormaux." L' Annee psychologique, XI, 1905. pp. 191-244.

14. ----------------- "Application des Methodes nouvelles au diagnostique du niveau intellectuel chez les enfants normaux et anormaux d'hospice et l'ecole primaire." L' Annee psychologique, XI, 1905. pp.245-336.

15. Binet, Alfred "Nouvelles recherches sur la mesure du niveau intellectuel chez les enfants d' ecole." L' Annee psychologique, XVII, 1911. pp. 145-201.

16. Binet, Alfred and Henri, V. "La psychologie individuelle." L' Annee psychologique, II, 1895. pp. 411-465.

17. Binet, Alfred and Simon, Th. "Le developpement de l'intelligence chez les enfants." L' Annee psychologique, XIV, 1908. pp. 1-94.

18. Brigham, Carl C. A Study of American Intelligence. Princeton University Press, 1923.

19. Broady, Knute O. School Provision for Individual Differences. Bureau of Publications, Teachers College, Columbia University, 1930.

20. Bronner, Augusta
Healy, William
Lowe, Gladys M.
Shimberg, Myra E.
Manual of Individual Tests and Testing. Little, Brown & Co.,1927.

21. Bruce, H. A. "Testing the Mental Tests." Century Magazine, CV, December, 1922. pp. 214-221.

22. Buckingham, Burdette R. Research for Teachers. Silver, Burdett & Co., 1926.

23. Cattell, J. McKeen "Mental Tests and Measurements." Mind, XV, 1890. pp. 373-380.

24. Cattell, J. McKeen and Ferrand, L. "Physical and Mental Measurements of the Students of Columbia University." Psychological Review, III, 1896. pp. 618-648.

25. Cattell, Psyche "Comparability of IQ's obtained from Different Tests at Different IQ Levels." School and Society, March 29, 1930. pp. 437-442.

26. Chapman, Harold B. Organized Research in Education. Bureau of Educational Research Monograph, #7. Ohio State University Press, 1927.

27. Cody, Frank "The Superintendent: His Administrative and Supervisory Staff." School Executives' Magazine, XLIX: 6, February, 1930. pp. 259-261.

28. Comings, W. R. "Are Mentality Tests on Right Lines?" Educational Review, LXIV, December, 1922. pp. 392-394.

29. Corning, H. M. After Testing -- What? Scott, Foresman Co., 1926.

30. Cunningham, Bess V. The Prognostic Value of a Primary Group Test. Bureau of Publications, Teachers College, Columbia University. 1923.

31. Davis, Helen The Use of Intelligence Tests in the Classification of Pupils in the Public Schools of Jackson, Michigan. Twenty-first Year Book of the National Society for the Study of Education, Chapter III. Public School Publishing Co. 1923.

32. Dickson, Virgil E. The Use of Mental Tests in School Administration. Board of Education, Berkeley, California, 1922.

33. ------------------ Mental Tests and the Classroom Teacher. World Book Co., 1924.

34. Dougherty, Mary L. A Comparative Study of Nine Group Tests of Intelligence for Primary Grades. Johns Hopkins Press, 1928.

35. English, H. B. "Is America Feeble-minded?" The Survey, October 15, 1922.

36. Fordyce, Charles "Intelligence Tests in Classification of Children in the Elementary Grades." Journal of Educational Research, IV, June, 1921.

37. Frazee, Laura "Administrative Aspects in Testing Individual Differences." Childhood Education, V. October, 1928.

38. Gesell, Arnold L. Mental Growth of the Pre-School Child. The Macmillan Co., 1925.

39. ------------------ Infancy and Human Growth. The Macmillan Co., 1928.

40. Goddard, Henry H. "A Revision of the Binet Scale." Training School Bulletin, VIII. Vineland, N. J., 1910.

41. ------------------ "Four Hundred Feeble-minded Children classified by the Binet Method." Pedagogical Seminary, XVII, 1910. pp. 387-397.

42. ------------------ "Two Thousand Children measured by the Binet Measuring Scale of Intelligence." Pedagogical Seminary, XVIII, 1911. pp. 232-259.

43. ------------------ Human Efficiency and Levels of Intelligence. Princeton University Press, 1920.

44. ------------------ School Training of Gifted Children. World Book Co.,1928.

45. Good, Carter V. How to do Research in Education. Warwick and York. Third Edition, 1929.

46. Goodenough, Florence L. Measurement of Intelligence by Drawings. World Book Co., 1926.

47. Haggerty, M. E. "Intelligence Examination, Delta 2." Journal of Educational Research, XIV, May, 1923.

48. ----------------- Haggerty Intelligence Examination. Manual of Directions for Delta 1 and Delta 2, Revision of 1928. World Book Co.

49. Huey, E. B. A Syllabus for the Clinical Examination of Children with the Revised Binet-Simon Scale for the Measurement of Intelligence. Warwick and York, 1912.

50. Hughes, W. Hardin Report from Bureau of Research and Service in Thirty-eighth Annual Report of the Pasadena Schools. 1924-'25.

51. Irwin, Elizabeth A. and Marks, Louis A. Fitting the School to the Child. The Macmillan Co., 1924.

52. Johnson, Eleanor H. "Mental Measurements." Mental Hygiene Bulletin, VII: 8, October, 1929.

53. Jones, Vernon A. Effect of Age and Experience on Tests of Intelligence. Bureau of Publications, Teachers College, Columbia University, 1926.

54. Kelley, Truman L. "Again Educational Determinism." Journal of Educational Research, VIII, June, 1923.

55. ----------------- Crossroads in the Mind of Man. Stanford University Press, 1928.

56. Keyes, C. H. Progress through the Grades of City Schools. Bureau of Publications, Teachers College, Columbia University, 1911.

57. Kohs, S. C. "The Binet Test and the Training of Teachers." Training School Bulletin, Vineland, N. J., X, 1914.

58. Kuhlmann, F. "Revision of Binet-Simon System for Measuring the Intelligence of Children." Journal of Psycho-Asthenics, Monograph Supplement, I: 1, September, 1912.

59. ------------------ A Handbook of Mental Tests. Warwick and York, 1922.

60. Layton, Warren K. The Group Intelligence Testing Program of the Detroit Public Schools. Twenty-first Year Book of the National Society for the Study of Education. Public School Publishing Co., 1923.

61. Lippmann, Walter Articles in the New Republic, Vols. XXXII and XXXIII.
(1) "Mental Age of Americans." - October 25, 1922. pp.213-215.
(2) "Mystery of the 'A' Men." - November 1, 1922. pp.246-248.
(3) "Reliability of Intelligence Tests." - November 8, 1922. pp. 275-277.
(4) "Abuse of the Tests." - November 8, 1922. pp.297-298.
(5) "Tests of Hereditary Intelligence." - November 22, 1922. pp. 328-330.
(6) "A Future for the Tests." - November 29, 1922.pp. 9-10.

62. Lundberg, L. D. "Tentative Scheme for Classifying Pupils into Homogeneous Groups." Mimeographed material from Flint, Michigan, School Department. 1928.

63. Maller, Julius B. Coöperation and Competition. Bureau of Publications, Teachers College, Columbia University, 1929.

64. McCall, William A. How to Measure in Education. The Macmillan Co., 1923.

65. ----------------- How to Experiment in Education. The Macmillan Co.,1923.

66. Mead, Margaret "Group Intelligence Tests and Linguistic Disability among Italian Children." School and Society, XXV. April 16, 1927.

67. Melville, N. J. Standard Method of Testing Juvenile Mentality by the Binet-Simon Scale. J.P.Lippincott Co., 1917.

68. Mort, Paul R. The Individual Pupil. American Book Co., 1928.

69. Nosworthy, Naomi "The Psychology of Mentally Deficient Children." Archives of Psychology, No. 1, November, 1906. Columbia Contribution to Philosophy and Psychology, XV: 2.

70. O'Brien, Francis P. The High School Failures. Bureau of Publications, Teachers College, Columbia University, 1919.

71. Orleans, Jacob S. A Study of the Nature of Difficulty. Bureau of Publications, Teachers College, Columbia University, 1926.

72. Pintner, Rudolph and Cunningham, Bess V. Manual of Directions for Pintner-Cunningham Primary Mental Test. World Book Co.,1923.

73. Pintner, Rudolph. "Training Students in Group Intelligence Testing." Journal of Educational Research, IX, 1924.

74. ----------------- "Pintner-Cunningham Primary Mental Test." Journal of Educational Psychology, XVIII: 1, January, 1927.

75. Reed, Mary M. An Investigation of Practices in First Grade Admission and Promotion. Bureau of Publications, Teachers College, Columbia University, 1927.

76. Root, W. T. "Correlations between Binet and Group Tests." Journal of Educational Psychology, XIII: 5, May, 1922. pp.286-292.

77. Ross, Clay Campbell The Relation between Grade School Record and High School Achievement. Bureau of Publications, Teachers College, Columbia University, 1925.

78. Sexson, John A. "Plan of Curriculum Administration and Research Investigation. Superintendent's Bulletin, No. 3, September, 1928. Pasadena City Schools.

79. Spearman, Charles E. Abilities of Man. The Macmillan Co., 1927.

80. ------------------ "Response to Kelley." Journal of Educational Psychology, XX: 8, November, 1929.

81. Stenquist, John L. "The Research Bureau meets Individual Differences." Childhood Education, V, October, 1928.

82. ------------------ The Research Bureau and the Schools. Baltimore Bulletin of Education, VIII: 1.

83. ------------------ Time and Labor-Saving Devices in Testing Programs. Mimeographed copy of Address (illustrated) given at School men's Week, University of Pennsylvania, March 22, 1928.

84. Teagarden, Florence M. A Study of the Upper Limits of the Development of Intelligence. Bureau of Publications. Teachers College, Columbia University, 1924.

85. Terman, Lewis M. Measurement of Intelligence. Houghton, Mifflin Co., 1916.

86. ----------------- The Intelligence of School Children. Houghton, Mifflin and Co. 1919.

87. ----------------- "The Psychological Determinist; or Democracy and the IQ." Journal of Educational Research, VI, June, 1922. pp.57-62.

88. ----------------- "The Great Conspiracy." The New Republic, XXXIII, December 27, 1922. pp. 116-120.

89. Terman, Lewis M. and Co-Workers. Genetic Studies of Genius, Vol. I. Stanford University Press, 1925.

90. Terman, Lewis M. "The Ultimate Influence of Standard Tests." Journal of Educational Research, XVII, January, 1928. pp. 57-59.

91. ----------------- Terman Group Test of Mental Ability. Manual of Directions.

92. Terman, Lewis M. et al. Stanford Revision and Extension of the Binet-Simon Scale for Measuring Intelligence. Warwick and York, 1917.

93. Terman, Lewis M. and Childs, H. G. "Tentative Revision and Extension of the Binet-Simon Scale." Journal of Educational Psychology, III, 1913.

94. Terman, Lewis M. et al. Intelligence Tests and School Reorganization. World Book Co., 1923.

95. Thompson, Thomas M. "Intelligence Tests and Democracy in Education." Educational Review, LXVII, January, 1924.

96. Thorndike, E. L. "The New Psychological Tests: Are they a substitute for Content Examinations?" Educational Review, LIX, 1920.

97. ----------------- "The Testing Movement in the Light of Recent Research." Journal of Educational Research, XVII, May, 1928. pp. 345-349.

98. Thorndike, E.L. et al. The Measurement of Intelligence. Bureau of Publications, Teachers College, Columbia University, 1927.

99. Thorndike, E. L., Terman, L. M. et al. "Intelligence and its Measurement." Journal of Educational Psychology, XII, March and April, 1921.

100. Torgerson, T. L. "Is Classification by Mental Tests and Intelligence Quotients Worth While?" Journal of Educational Research, XIII, March, 1926. pp. 171-180.

101. ----------------- "Tests published in 1928." Journal of Educational Research, XIX, May, 1923. pp. 376-381.

102. Trabue, Marion Rex. Measuring Results in Education. American Book Co., 1924.

103. Vincent, Leona. A Study of Intelligence Test Elements. Bureau of Publications, Teachers College, Columbia University, 1924.

104. Washburne, Carleton. "Selection of Under-age Children for Entrance into School." Educational Administration and Supervision, XIV: 3, March, 1928.

105. Wentworth, Mary M. Individual Differences in the Intelligence of School Children. Harvard University Press, 1926.

106. Whipple, Guy M. "Educational Determinism." School and Society, XV. June 3, 1922. pp. 599-602.

107. ----------------- "The Intelligence Testing Program and its objectors, conscientious and otherwise." School and Society, XVII, May 26, 1923, pp. 561-568, and June 2, especially pp.602-604.

108. ----------------- "Intelligence Tests and Their Use." Twenty-first Year Book of the National Society for the Study of Education, Part I.

109. ---------------- "Intelligence Tests in Colleges and Universities." Twenty-first Year Book of N. S. S. E .

110. Wilson, Guy M. "The Puspose of a Standardized Test in Spelling." Journal of Educational Research, XX: 5, December, 1929.

111. Wilson, Guy M. and Hoke, Kremer J. How to Measure. Second edition. The Macmillan Co., 1928.

112. Worlton, J. T. "The Why of Homogeneous Classification." Elementary School Journal, XXVII, December, 1926. pp. 265-274.

113. Yerkes, Robert M. Psychological Examining in the United States Army. Memoirs, National Academy of Sciences, XV, 1921. p. 785 and ff.

114. Yerkes, Robert M.)
Bridges, James W.)
Hardwick, Rose S.) A Point Scale for Measuring Mental Ability. Warwick and York, 1915.

115. National Research Council, Manual for National Intelligence Tests. World Book Co.

116. Boston Post, April 9, 1922.

117. Bulletin #67. Board of Education, Cleveland Bureau of Educational Research, April 8, 1927. File No. 1-02-04. (Mimeographed)

118. School Executives Magazine, February, 1930.

AUTOBIOGRAPHY OF NELLIE EVA POWERS

Place of birth,	Westborough, Massachusetts
Date of birth,	February 24, 1877
Father,	Frank Waldo Powers of Westborough, Massachusetts
Mother,	Calista Ida Achorn-Powers of Waldoborough, Maine

Schools and Colleges attended:

1895	Graduate, High School, Westborough, Mass.
1898	Graduate, State Normal School, Worcester, Massachusetts
1903,'05,'26	Special Student, Clark University
1911	Graduate, Training School, Vineland, N. J. (Included observation work at University of Pennsylvania)
1912	Special Student, New York University
1913	Special Student, with Dr. Arnold L. Gesell of Yale University
1921 - '30	Special, Undergraduate, and Graduate Work, Boston University
1928	Graduate Student, Columbia University

Degrees:

B.S. in Education, Boston University, 1923
A.M. , Boston University, 1925

Scholastic positions filled:

1898-1902	Principal, Center Grammar School, Auburn, Mass.
1902-1903	Substitute in all grades up to ten, Worcester City Schools
1903-1908	Instructor, Special Classes for Defectives, Worcester
1908-1913	Instructor, Special Class, Boston
1912	Instructor, Dept. of Abnormal Psychology Summer Session, New York University
1913-----	In Charge at Frances E. Willard School, Boston
1926-1929	Organizer of Measurement in Public Schools, Belmont, Mass.
1927,1929	Instructor, Measurement of Intelligence, Summer Session, Boston University

BOSTON UNIVERSITY
1 1719 02572 9429